

RESEARCH TRIANGLE INSTITUTE

RTI/1796/00-03F

NASA CR-165735

CONCEPTS FOR ON-BOARD SATELLITE IMAGE REGISTRATION

Final Report

Volume Three

**Impact of VLSI/VHSIC on Satellite
On-Board Signal Processing**

Prepared for



National Aeronautics and Space Administration
Langley Research Center
Hampton, Virginia

CONCEPTS FOR ON-BOARD SATELLITE IMAGE REGISTRATION

Final Report

Volume Three

IMPACT OF VLSI/VHSIC ON SATELLITE ON-BOARD SIGNAL PROCESSING

Prepared Under Contract NAS1-15768

by

J V Aanstoos, W H Ruedger

Research Triangle Institute
Research Triangle Park, North Carolina 27709

and

W E Snyder

North Carolina State University
Raleigh, North Carolina 27650

Prepared For



National Aeronautics and Space Administration
Langley Research Center
Hampton, Virginia

July 1981

PREFACE

This report was prepared by the Research Triangle Institute, Research Triangle Park, North Carolina, under Contract NAS1-15768. The work has been administered by the Electronics Devices Research Branch of the Flight Electronics Division, Langley Research Center, National Aeronautics and Space Administration. Mr. W. L. Kelly IV served as Technical Representative.

These studies began on 23 May 1980 and were completed on 15 December 1980. Mr. W. H. Ruedger served as Project Leader. Mr. J. V. Aanstoos completed the project team. Dr. W. E. Snyder, North Carolina State University, served as Consultant to the program.

TABLE OF CONTENTS

1 0	Introduction	1
2 0	DOD Program Description	2
3 0	Technology Forecast	4
3 1	Scaling of LSI Systems	4
3 2	Speed	8
3 3	Chip Density	12
3 4	Reliability	14
3 5	Radiation Tolerance	18
3 6	Two Sample Designs	20
4 0	On-Board Signal Processing Functions	35
4 1	Convolution and Correlation	35
4 2	Transforms	42
5 0	VHSIC Availability	48
6 0	Beyond VHSIC	49
	References	54

1 0 Introduction

NASA has embarked on a program to increase the effectiveness and efficiency of the system that couples the user of space data with the sensors that acquire this data. This program, the NASA End-to-End Data System (NEEDS), addresses the identification, development, and demonstration of data handling techniques and technologies which are required to accomplish these goals.

More specifically, the NEEDS program goals present a requirement for on-board signal processing to achieve user-compatible, information-adaptive data acquisition. These signal processing functions comprise a major constituent of the Information Adaptive System (IAS), a significant module of the NEEDS concept. The IAS essentially consists of the spaceborne portion of NEEDS exclusive of telemetry, support, and housekeeping functions.

This volume addresses the impact of anticipated advances in microelectronics technology on on-board signal processing systems, as evidenced by the Defense Department's Very High Speed Integrated Circuits (VHSIC) program which is described in Section 2 0. Section 3 0 presents a technology forecast, with predictions of improvements in speed, density, power consumption, and reliability. A discussion of the radiation tolerance of the new technology and sample designs are also included in this section. Section 4 0 discusses important on-board signal processing functions and their implementations, and how they will be affected by the new technology. Section 5 0 forecasts availability of systems implemented using VLSI, and Section 6 0 looks beyond the VHSIC program to future VLSI improvements, particularly the less mature Gallium Arsenide technology.

2 0 DOD VHSIC Program [1] [2]

The Department of Defense has initiated the VHSIC (Very High Speed Integrated Circuits) research and development program for the following expressed purposes

- To obtain high throughput signal and data processing capability for military systems
- To accomplish life-cycle cost reductions
- To assure complex IC capabilities in military electronics
- To provide affordable systems

The program has several specific goals for the advancement of IC technology. The major goals are

- Function throughput rate increase from the current 1.5×10^{11} gate-hz/cm² to 5×10^{11} gate-hz/cm² initially and ultimately to 10^{13} gate-hz/cm²
- Easy insertion of new technology
- Radiation tolerance
- Availability for application in any military system
- Built-in test at the chip level

The VHSIC program is structured in three serial phases and one concurrent phase. Phase zero, recently completed, was a nine-month concept definition period for developing approaches to system architecture, chip architecture and design, IC processing technology, and testing. The three-year Phase I will bring 1.25 micrometer IC's into pilot production, subsystem brassboards will be designed and developed and sub-micrometer IC development will begin. In the final 30-month Phase II, the 1.25 micrometer brassboards will be demonstrated and the sub-micrometer IC's will be brought into pilot production. Along with this three phase main-stream effort is the concurrent Phase III technology support effort. The central purpose of this phase is to provide a broad base of technical support to the main program with emphasis on innovation through limited-scope programs that focus on key technologies, sub-systems, processes, equipment, architecture, and computer-aided-design tools and techniques. Figure 2-1 shows the proposed time-frame of these milestones.

The DOD VHSIC program is confined to silicon technologies I^2L , NMOS, CMOS, CMOS-SOS, and the variants of these which are being investigated. Research into other technologies such as gallium arsenide or Josephson Junction IC's which are oriented to post-1987 applications is being supported by DARPA and the individual services.

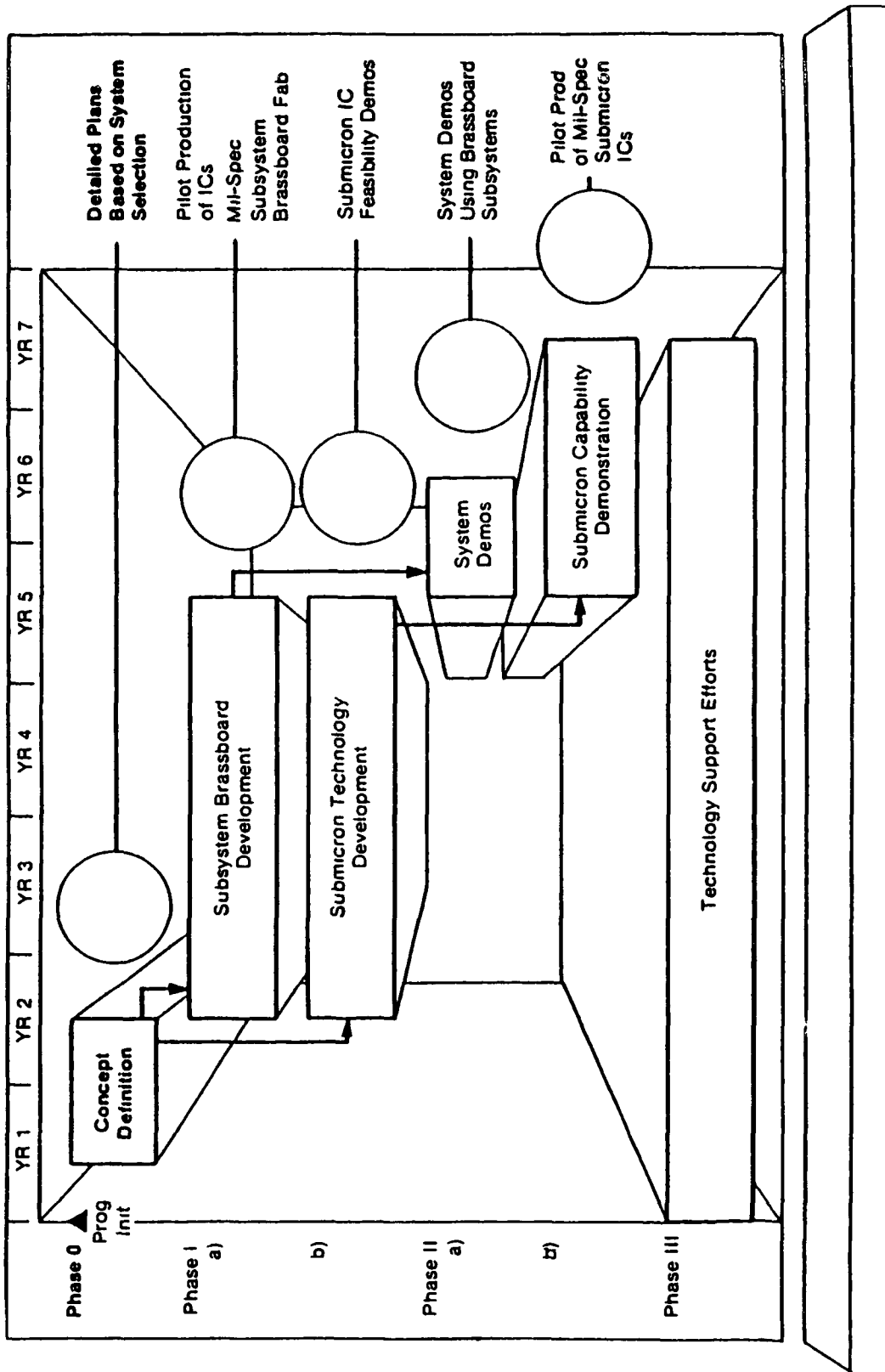


Figure 2-1 Major Milestones for DOD VHSIC Program [1]

3 0 Technology Forecast

VHSIC is a DOD initiative to encourage development of VLSI to meet military reliability and service requirements

VHSIC III specifically addresses silicon. It is short range in the sense that each activity should provide usable products - either hardware, software, or knowledge - "that can be incorporated in brassboard demonstrations of VHSIC technology or contributed to the design and pilot production of 1.25 μm design-rule ICs and then sub- μm design-rule ICs within the seven-year span of the program" [1]. This requirement for demonstrability effectively confines VHSIC to silicon technologies: bipolar (including I^2L and its variants), NMOS, CMOS, and CMOS-SOS.

Gallium-arsenide gates currently are at least five times faster than silicon gates, but the GaAs digital IC technology is far less mature than silicon digital technology. The primary obstacle to this development has been the lack of a reliable oxide-insulated gate for GaAs MESFET.

Although GaAs technology is not included in the VHSIC program, and consequently is not discussed further in this section, industry observers predict that

GaAs digital VLSI will be achieved by 1985 [2] (a figure that may be somewhat optimistic).

GaAs digital LSI will not compete directly with silicon chips for the same applications, but GaAs ICs will complement silicon ICs for gigabit applications beyond the capability of silicon.

3 1 Scaling of LSI Systems

The goal of the VHSIC program is pilot production in 1986 of chips containing 250,000 gates operating at clock speeds of 25 MHz. These gates would be fabricated by MOS or bipolar technology and have minimum dimensions of 5 to 8 μm . The required speed and circuit density would be obtained both by scaling down current LSI circuits (reducing channel length, oxide thickness, and supply voltage) and by developing new types of system architecture and software.

3 1 1 Scaling Rules

To date, three technologies have emerged that are reasonably high in den-

sity and scale to submicron dimensions without an explosion in power per unit area [3] These are the n-channel MOS silicon gate process (NMOS), the complementary MOS silicon gate process (CMOS), and the integrated injection logic (I^2L) process

3 1 2 Scaling of NMOS

As NMOS is one of the most popular candidate technologies for VLSI at present, a good deal of information is available on the scaling of parameters For example, as of 1978, typical MOS electrical parameters were

Resistances per Square

Metal	=	0.03	Ω / \square
Diffusion	=	10	Ω / \square
Polysilicon	=	15-100	Ω / \square
Transistor Channel	=	10,000	Ω / \square

Capacitances

Gate-channel	$4 \times 10^{-4} \text{ pf}/\mu\text{m}^2$
Diffusion	$1 \times 10^{-4} \text{ pf}/\mu\text{m}^2$
Polysilicon	$0.4 \times 10^{-4} \text{ pf}/\mu\text{m}^2$
Metal	$0.3 \times 10^{-4} \text{ pf}/\mu\text{m}^2$

Consider the MOSFET shown in Figure 3-1 Scaling this transistor's dimensions and gate voltage V (the gate-source Voltage V_{gs} minus the threshold voltage of the transistor, V_{th}) by a factor α (such that $L' = L/\alpha$, $W' = W/\alpha$, $D' = D/\alpha$, and $V' = V/\alpha$) causes the resistances per square to scale up by α (except

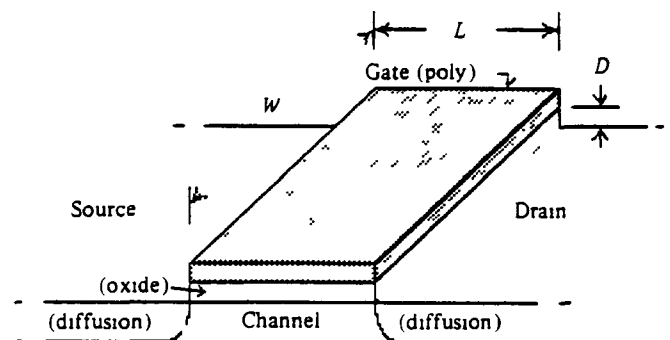


Figure 3-1 MOSFET Construction [3]

transistor channel resistance, which will be independent of α), and the capacitances to scale up by α

The resulting transistor's parameters scale as

Transit time $r' = r/\alpha$

gate capacitance $C' = C/\alpha$

drain to source current $I' = I/\alpha$

Both switching power and static power per device scale down by $1/\alpha^2$

The switching energy per device (defined as the power consumed at maximum clock frequency multiplied by device delay) scales as

$$E'_{sw} = E_{sw}/\alpha^3$$

An operational parameter which may be derived from these parameters is the functional throughput rate (FTR), gates per chip multiplied by clock speed per gate. In NMOS, FTR scales by α^3

If one extends the possible scaling of MOS to the limits imposed by physical law [4], one can see the potential scaling offers

	1978	19XX
Minimum feature size	6 μ m	0.3 μ m
r	0.2 to 1 ns	0.2 ns
E_{sw}	10^{-12} joule	2×10^{-16} joule
System clock	30 to 50 ns	2 to 4 ns

The limit of 0.3 μ m on channel length is due to the fact that at that point, physical effects such as tunneling through the gate oxide and fluctuations in dopant densities in the depletion layers make smaller devices unworkable

Thus, scaling down an integrated system built with NMOS technology by a scale factor of $\alpha=10$ will produce a system having one hundred times the circuits per unit area. The total power per unit area remains constant. All voltages are reduced by a factor of 10, and therefore the current supplied per unit area increases by a factor of 10. The time delay per stage is decreased by a factor of 10. Therefore, the power-delay product decreases by a factor of 1000.

The increase in current density causes a limitation to scaling other than the 0.3 μ m limit mentioned earlier. A current flux exceeding a certain limit (10^5 A/cm² in Al) through a metal conductor causes the metal atoms to move slowly in

the direction of the current. If there is a small constriction in the metal, the current density will be higher at that point, and more metal atoms will be carried forward from that point, narrowing the constriction still more.

The 10^5 A/cm^2 converts to a few milliamps per square micron, a value currently approached in present systems. Thus, metal thickness cannot be permitted to scale in the same way as other dimensions do.

In addition, short pulses of current seem much less prone to causing metal migration than DC currents, a factor which seems to favor processes like CMOS that do not require static DC currents.

3.1.3 Scaling of Other MOS Technologies

Any technology in which a capacitive layer on the surface induces a charge to flow under it to form a voltage-controlled transistor will scale in the same way that NMOS scales. Such technologies include MESFET's, Junction FET's, and CMOS devices.

Vertical MOS (VMOS) and Double-diffused MOS (DMOS) are both attempts to build MOS-type transistors, but to make use of controlled doping profiles to achieve extremely narrow channel widths. At ultimately small dimensions, silicon-gate processing should be able to achieve comparable channel lengths with simpler processing steps. Conway and Mead [3] state that these two technologies "while competitive at present feature size, are likely to be interim technologies that will present no particular advantage at submicron feature sizes." This statement is probably true for large scale digital systems, which is the emphasis of their book, however, VMOS in particular has applications to high speed-high power switching which is likely to keep it an active technology for those applications.

3.1.4 Scaling of Bipolar Technologies

The term "bipolar" refers to the fact that both types of carriers, holes and electrons, are involved in the operation of the device, whereas in MOS type devices only one carrier (electrons in NMOS) is involved. To be proper then, one really should distinguish between "vertical bipolar devices" such as NPN transistors, and "planar bipolar devices" such as the lateral transistors used in I^2L .

Traditionally, vertical bipolar circuits have been fast because their transit time was determined by the extremely narrow base width of the devices. The

base widths of high performance vertical bipolar devices are already nearly as thin as device physics allows. For this reason, the delay times of vertical bipolar devices are expected to remain approximately constant as their surface dimensions are scaled down. As both technologies approach their physical limits, the base width of vertical bipolar devices and the channel length of FET devices are limited by the same basic set of physical constraints and are therefore similar in dimension.

No single set of scaling principles applies to all bipolar gates, since the devices are more varied and complex than MOS gates. However, the scaling of voltages is generally inapplicable to bipolar technologies, including I^2L , since supply voltages are already at the physical minimum and constant-voltage scaling must be used. The power-delay product then scales as $1/\alpha$ instead of $1/\alpha^3$.

3.2 Speed

Most published studies and forecasts report the power-delay product rather than simply the speed. This is most obviously required in the case of I^2L , where one can simply inject a higher current from the power supply and achieve higher speed at the cost of higher power dissipation. The graph in Figure 3-2 represents the results of detailed simulations and available data relating the size of the device to the power-delay product.

Ferranti [5] breaks the current technologies down in greater detail, spelling out state-of-the-art speeds for devices which exist today, but may be in the laboratory stage. This information is included here as Tables 3-1, 3-2, and 3-3.

Table 3-1 General Performance Characteristics of I^2L Technologies [5]

TECHNOLOGY PARAMETER	I^2L	C^3L	CHIL	SFL	Schottky I^2L	SCTL	STL	I^3L	TTL
Propagation Delay (min) nS	10	3	32	8.8	5	5	5	4	10
Speed-power pJ/gate	1.0	3	0.25	0.07	0.5	0.2	0.2	0.6	10
Gates/mm ²	300	150	200	650	250	NOT QUOTED	NOT QUOTED	200	20
Logic Voltage Swing (volts)	0.7	0.25	0.3	NOT QUOTED	NOT QUOTED	0.3	0.15	NOT QUOTED	5

Detailed simulations and available data confirm the major predictions of scaling theory applied to MOS and bipolar (I^2L) gates. The curves indicate the scaling relationships for power delay product for gates with a fanout of 1. The dotted line indicates the approximate scaling of power delay with the minimum dimension, d .

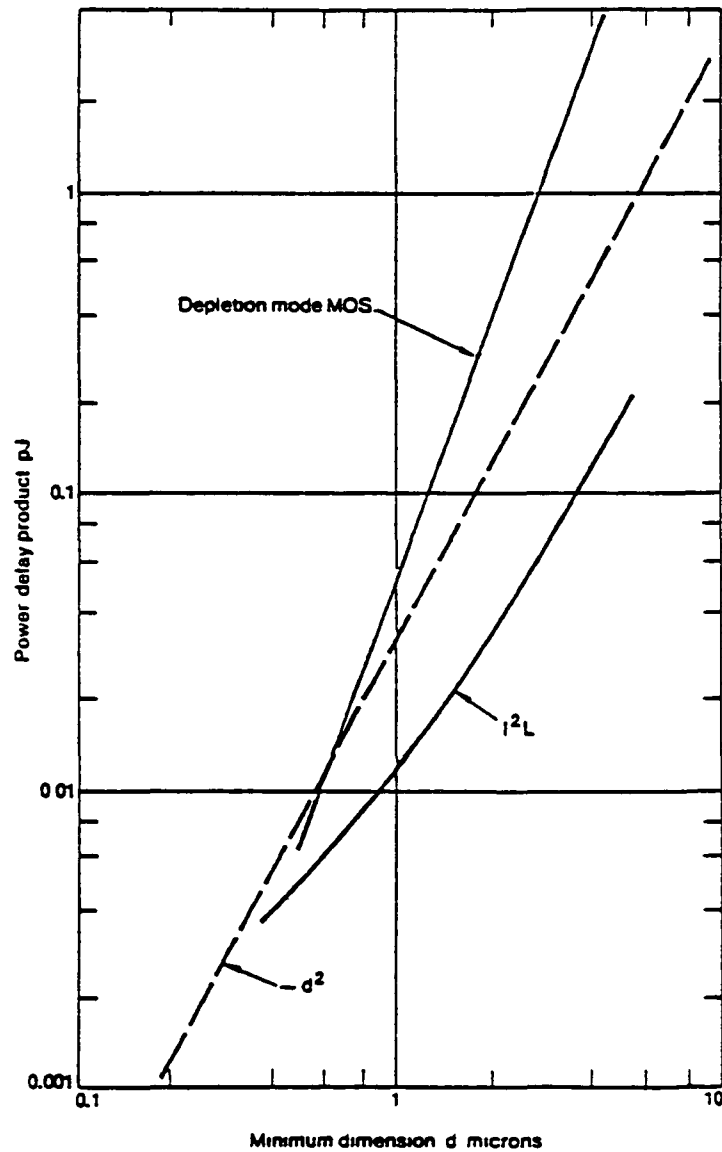


Figure 3-2 Relation Between Device Size and Speed-Power Product [2]

Table 3-2 General Characteristics of ECL [5]

FEATURE	MECL I	MECL II	MECL 10,000		MECL III
			10,100 Series 10,500 Series	10,200 Series 10,600 Series	
Gate Propagation Delay	8 ns	4 ns	2 ns	1.5 ns	1 ns
Gate Edge Speed	8.5 ns	4 ns	3.5 ns	2.5 ns	1 ns
Flip-Flop Toggle Speed (min)	30 MHz	70 MHz	125 MHz	200 MHz	500 MHz
Gate Power	31 mW	22 mW	25 mW	25 mW	60 mW
Speed-Power Product	250 pJ	88 pJ	50 pJ	37 pJ	60 pJ
Transmission Line Capability	No	On Some Devices	Yes	Yes	Yes
Output Pulldown Resistors	Yes	Optional	No	No	No
Input Pulldown Resistors	No	No	50 K Ω	50 K Ω	2K Ω & 50 K Ω

Table 3-3 Comparison of Basic MOS Gate Performance [5]

TECHNOLOGY	Propagation Delay (nS)	Speed-power product (pJ)	Chip Density		Typical Chip Size (mm^2)
			Devices/ mm^2 (1)	Gates/ mm^2	
High threshold P-channel metal gate	80	450	150	50	7 x 7
P-channel silicon gate	30	145	270	90	6.5x6.5
N-channel silicon gate	15	45	285	95	6 x 6
N-channel silicon gate depletion load(2)	12	38	320	107	6 x 6
N-channel double poly(3)	10	35	525	175	6 x 6
CMOS silicon gate	10	0.5	220	45	5.5x5.5
CMOS/SOS	2-5	0.005	650	275	5 x 5
VMOS DMOS	5	20	600	225	Not quoted

(1) 'Devices' in this context means transistors

(2) Depletion load is an MOS transistor used as a high impedance resistive load by connecting its gate electrode and drain

(3) a technique used to reduce cell size and hence reduce parasitic capacitance. Mainly used in memory devices.

Leonard [6] compares the current and future technologies as follows

	Propagation Delay (ns)	
	1980	1985
ECL	0.6	0.3
TTL	2	1.5
HMOS (NMOS)	5	1
CMOS	6	1
I^2L	3	1
ISL	3	0.4

This table may be used in conjunction with the extrapolated sizes of devices to obtain results which compare favorably with Sumney's objective [2] for VHSIC of several million to several billion operations per second

3.3 Chip Density

The goal of the VHSIC program is pilot production in 1986 of devices containing 250,000 gates [2]. Other often-quoted numbers are an ultimate clock rate of 25 MHz and a functional throughput of 10^{13} gate-hertz. To put these numbers into some sort of meaningful relationship requires a definition of the word "gate". This is a difficult term, since the way one produces a gate may be quite different in different technologies as, say I^2L and NMOS.

As a working definition in order to get a feel for the area, one may define a "gate" as an inverter followed by a pass transistor and a contact cut. These concepts are very germane and meaningful in the context of NMOS. In that technology, an inverter followed by a pass transistor is a fundamental building block, most obviously used in shift registers, but also an inherent part of many other circuits. In I^2L , however, pass transistors are not used, and consequently this example is not appropriate to I^2L .

Figure 3-3 below shows a physical layout of a typical implementation of this circuit. It is interesting to note that the contact cut required to connect the diffused region to the polysilicon conductor actually requires about five times the area of the pass transistor. Thus while counting active devices provides one means of measuring the complexity of a circuit, such a measure can often be somewhat misleading. There are other ways to lay this device out which increase the

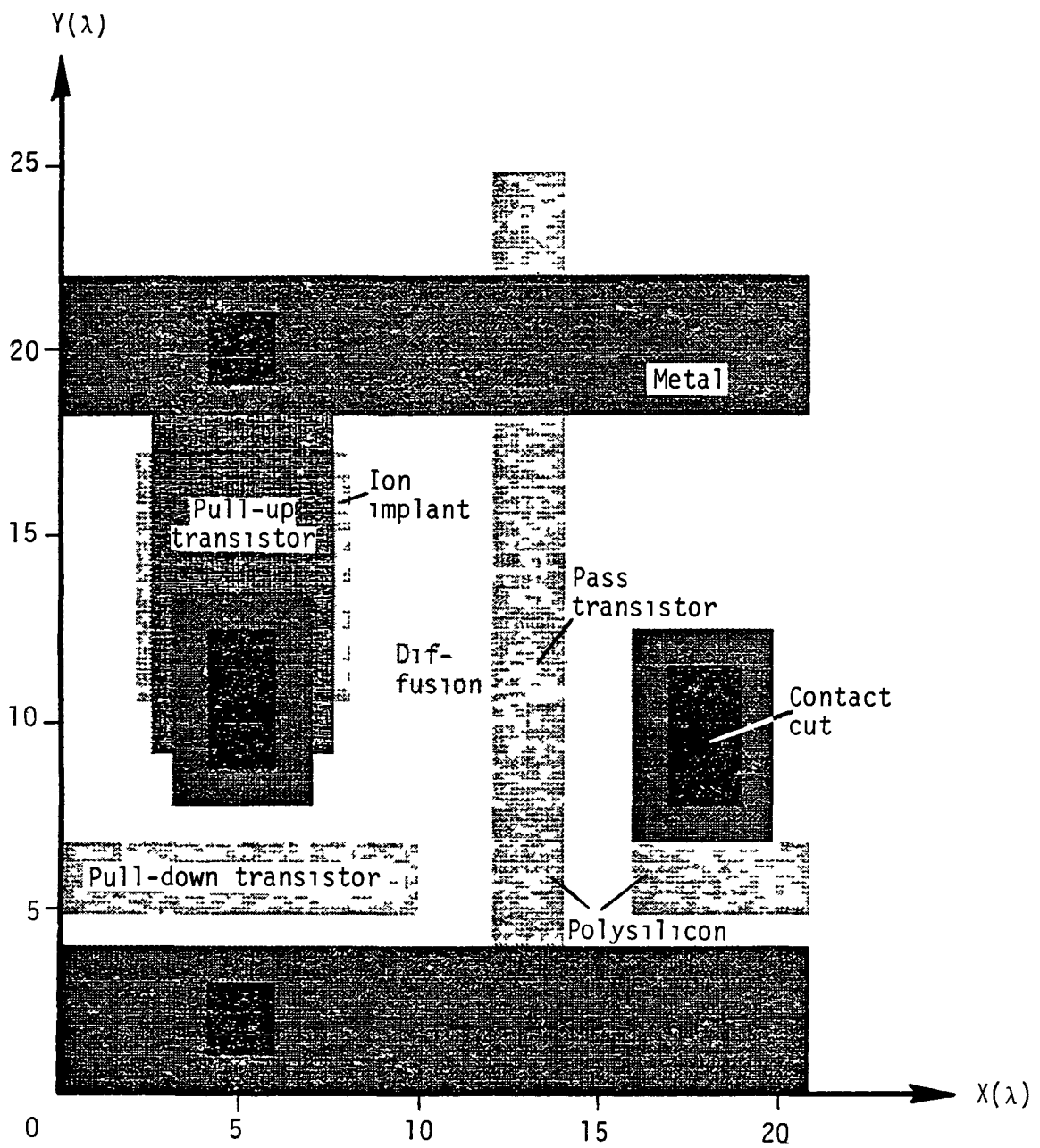


Figure 3-3 Typical Gate Layout [3]

density slightly, however, this will do for illustration

The gate requires a chip area of $25\lambda \times 21\lambda$ or $525\lambda^2$. Assuming a one micron technology, this converts to an area per gate of $5.25 \times 10^{-10} \text{ m}^2$. Thus, one could pack 1.9×10^5 of these per square centimeter.

At this packing density, Sumney's FTR of 10^{13} gate-hertz/cm² will require clock rates of 52MHz. At the upper limit of technology, $0.3\mu\text{m}$, the density goes up to 2.1×10^6 gates per cm², and the clock rates needed to achieve 10^{13} g-hz/cm² reduces to 4.7MHz.

One interesting contradiction is that with smaller devices, lower clock rates are needed to achieve the desired FTR, yet with smaller devices, it is actually easier to achieve higher clock rates. Consequently, FTR available from a technology can be said (loosely of course) to vary as the square of the "smallness" achievable by that technology.

3.4 Reliability [7]

There are several factors influencing the reliability and maintainability of current and future digital systems.

The first of these factors is simply the improvement in fabrication technology. As the semiconductor manufacturers develop better mask alignment procedures, better photography, improved control over chemistry, etc., the reliability of the products produced increases. Figure 3-4 shows the decrease in failure rate for the same part, the Motorola 6800 microprocessor, comparing the failure rate for parts manufactured in 1975, up to 1979. The curve is also extrapolated to the predicted VLSI 32 bit microprocessor.

The second factor impacting reliability is circuit density. As gates per chip increase as shown in Figure 3-4, the failure rate per gate drops off almost as the inverse of this curve. This can be attributed to several factors.

System failures are often approximately constant per chip, since most failure mechanisms have to do with interconnects. Consequently, increasing gates per chip decreases failures per gate and improves system reliability. Higher density chips also must be designed for less power dissipation per gate. Lower power means lower failure rate, thus compensating for the increased number of components and making the "constant failure rate per chip" assumption close to true.

The combined impact of these factors can be seen in Figure 3-5, comparing qualitatively the reliability as a function of time for several technologies.

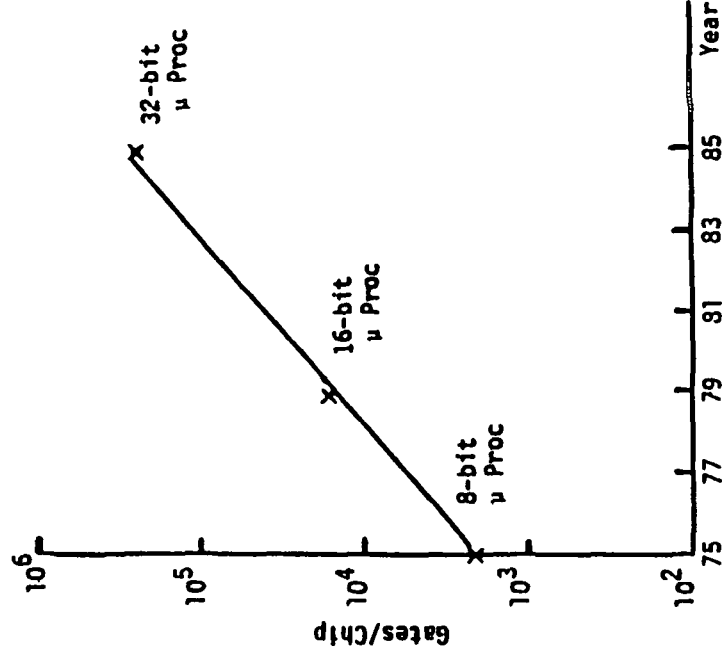


Figure 3-4b Microprocessor Circuit Density Increase [7]

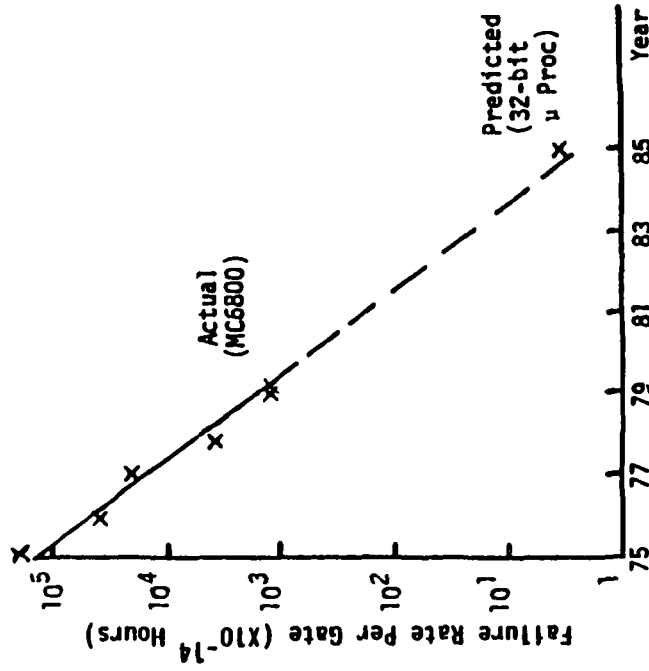


Figure 3-4a Microprocessor Failure Rates [7]

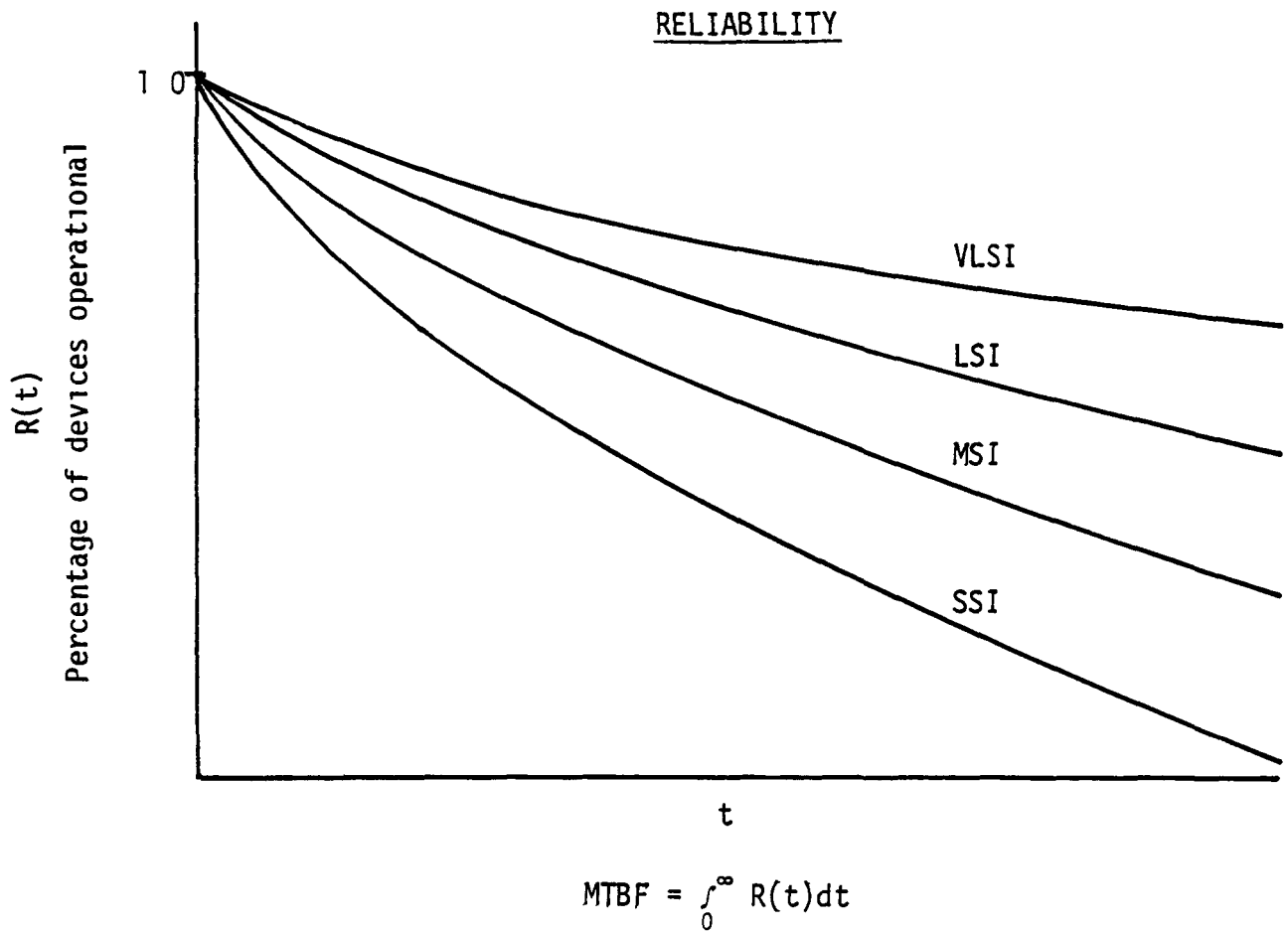


Figure 3-5 Reliability as a Function of Time

- Increased gates per chip reduces failure rates directly by
 - reducing number of interconnections
 - reducing power requirements
 - reducing power dissipation per gate
- Reliability can be further enhanced through use of redundancy-based fault tolerant architectures made possible by increased chip complexity and reduced cost per gate

The horizontal time axis is scaled in thousands of operational hours. The MTBF is the integral under the appropriate curve.

3.4.1 Alternative Maintenance Approaches Made Possible by New Technology

To a large extent, system reliability and maintainability can be achieved through the intelligent incorporation of redundancy. Careful redundancy management is often the key to achieving reliable operation, as well as maintainable systems.

Redundancy implies added costs. Since the cost of integrated circuits on a "per gate" basis is dramatically decreasing, advantage must be taken of these happy circumstances.

3.4.2 Reliability Enhancement

There are two fundamental approaches to achieving high system reliability. These approaches are to 1) use high reliability components and 2) use redundant system resources. The first technique has been described as fault intolerant, while the second method is referred to as a fault tolerant method [8]. The first method, which involves the use of high reliability components, while straightforward in concept, is expensive in practice. Existing military maintenance approaches make use of this concept.

The achievement of enhanced reliability through the use of fault-tolerant computing methods is becoming more popular as high-density integrated circuits become available. Much work has been done in recent years in the area of fault-tolerant computing for space program applications [9]. However, the methods which are used to achieve fault tolerance in spaceborne systems can be applied to military ship-board, ground-based and airborne applications, provided there are maintenance concepts in place to accommodate such approaches.

Fault tolerance can be achieved through the use of various forms of protective redundancy. Such methods should be applied to different systems with attention given to factors such as the types of faults, performance, life-cycle, etc. Basically, protective redundancy can be introduced in three different forms:

1. additional hardware
2. additional software
3. repetition of operations

Particular fault-tolerant approaches which make use of such redundancy are de-

scribed below

Static redundancy - In static hardware redundancy, faults are masked by the use of additional modules. The most common and practical types of static redundancy are replication of components, Triple Modular Redundancy (TMR), and N-Modular Redundancy (NMR).

Dynamic redundancy - This deals with two types of modules. One of these contributes directly to the output of the system (active modules), and the others are the standby spares which replace a failed module. In dynamic-redundant systems the fault-tolerant operation is realized by three consecutive actions: fault detection, fault diagnostic, and recovery. Usually, other methods of protective redundancy, such as software and time redundancy, are employed in dynamic redundancy.

Hybrid redundancy - This is a combination of static and dynamic redundancy, in other words, the standby spare and the active modules themselves use static redundancy.

In software-redundant systems extra software routines are added to the system. In time-redundant systems faulted operations are repeated several times.

Figure 3-6 is an example of the trade-offs that can be made between reliability enhancement through the use of high reliability components and through the use of modular redundancy. This example illustrates the reliability function $R(t)$ for a Digital Equipment Corporation LSI-11. The failure rate (λ) of the LSI-11 was calculated using a parts count model based on MIL-STD-217B. The resulting reliability function is $R(t)$ in Figure 3-6. A corresponding reliability function, $R(t)_{TMR}$, for a triple modular redundant LSI-11 with an ideal voter is also given for the predicted LSI-11 failure rate, λ . The third curve in Figure 3-6, $R'(t)$, represents the reliability function for a Digital Equipment Corporation LSI-11, with assumed highly reliable components with failure rate of $\lambda' = 0.1\lambda$.

3.5 Radiation Tolerance [10]

The exact mechanisms for device failure in radiation environments are not well known. Several effects can be hypothesized with reasonable certainty however.

MOS devices are most sensitive to radiation, due to charge trapping in the dielectric. Self-aligning gate structures are the worst, since the process pre-

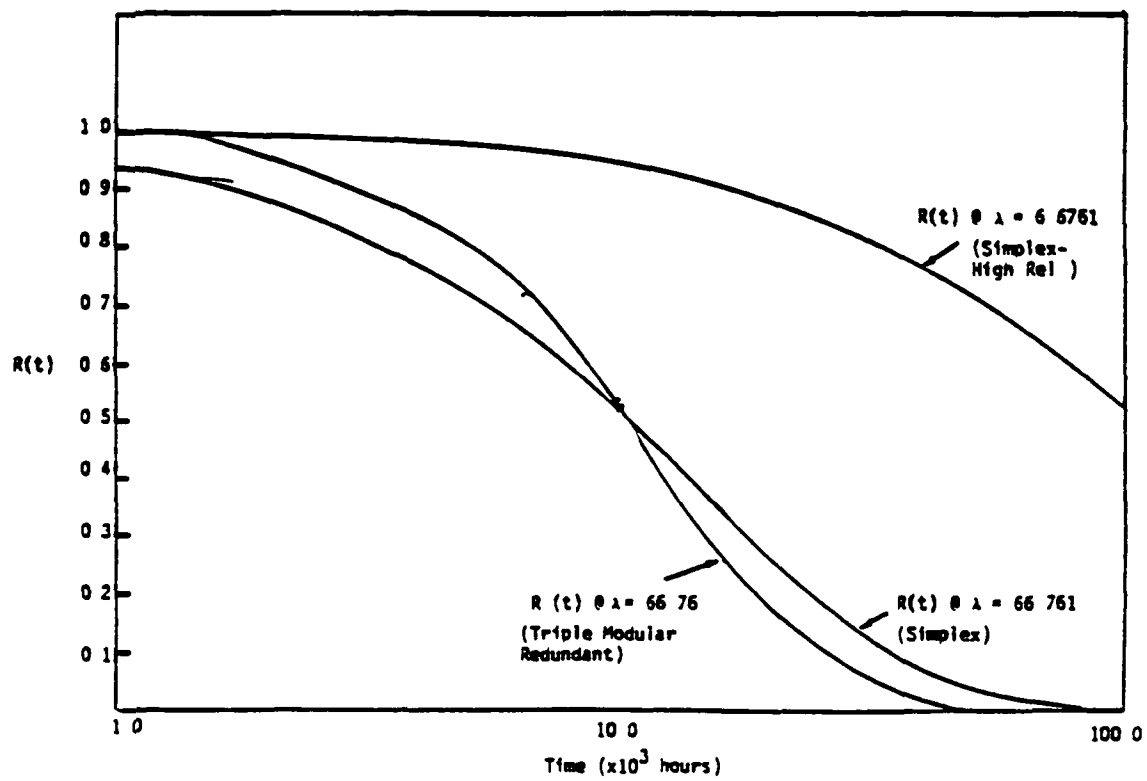


Figure 3-6 Implementation Alternatives for Reliable LSI-11 [7]

vents stripping and regrowth of the gate oxide late in fabrication. Specifically, structures to avoid are

- Two-level gate structures

- Polysilicon-gate structures

- Surface-channels (i.e., use buried channels)

By comparison with the MOS device, the bipolar transistor is only mildly affected by space radiation of the levels in question here. However, the minority carrier devices are still more susceptible than devices such as silicon JFETs and Gallium Arsenide MESFET's. There does not appear to be significant difference between the hardness of these two devices, however, there has been little integrated circuit technology developed around the silicon JFET since there are easier ways to accomplish the same (non-hardened) functions in silicon. In GaAs, however, since MOS devices do not exist, a significant amount of work has been done toward the development of integrated circuits.

In a high Gamma flux transient environment, the GaAs device will also perform better than a silicon JFET due to the direct band gap of the material.

These factors probably explain why DARPA has chosen GaAs as the material for the Advanced On-Board Signal Processor.

3.6 Two Sample Designs

In this section, two on-board signal processing functions are examined: 1) radiometric correction and 2) along-scan geometric correction. Two techniques for designing systems to perform these tasks are examined, using special-purpose, dedicated hardware, and using a microprogrammed central processing unit.

In the case of radiometric correction, it is immediately shown that special-purpose hardware is a more effective approach to the design than use of a CPU.

In the case of the along-scan processor, a fairly detailed design is undertaken utilizing the AMD 2900 chip set. RTI's philosophy in performing this design is that the 2900 family represents the current state-of-the-art in micro-programmable processors, and therefore represents one of the more attractive mechanisms for implementing special-purpose, high-speed processing functions. The assumption is that with the advent of VHSIC, it will be possible to integrate all the functions of the 2900 chip set onto a single chip.

The design of the along-scan processor is carried to a device count detail, and some alternative architectures are discussed.

As a result of this study, it is observed that the general purpose CPU architecture, even though it is microprogrammable, is not well suited for this on-board task either. However, the microprogram sequencer (such as the 2909) is a useful control component of a special purpose signal processing system.

3.6.1 Radiometric Correction

Assuming a line scanner 6000 elements long, it is necessary to correct the output of each cell by first subtracting an offset due to dark current and then multiplying by a scale factor. It is infeasible to perform such operations using a general purpose, or even a microprogrammed machine at the data rates needed. The calculations are as follows:

For 30 m accuracy, at nominal satellite velocity, the line scanner must be read every 4.44 ms. With 6000 pixels per line, we require a pixel to be read every 740 ns. Assuming we process seven spectral bands, a pixel must be corrected every 105 ns.

Assuming a high degree of integration, one could configure a bipolar processor on a chip. The microcycles for this processor would be:

- 1 read data, pixel number, band number
- 2 look up offset and scale factor
- 3 subtract offset from data
- 4 multiply result by scale factor
- 5 store output

Five microcycles are needed, even if one assumes sufficient parallelism to perform the read operation (1 above) in one microcycle. To build a chip with such parallelism is to produce a special purpose device rather than a general purpose processor. If one concedes the need for special purpose processing, then special purpose hardware can be built which is much simpler than a microprogrammable machine.

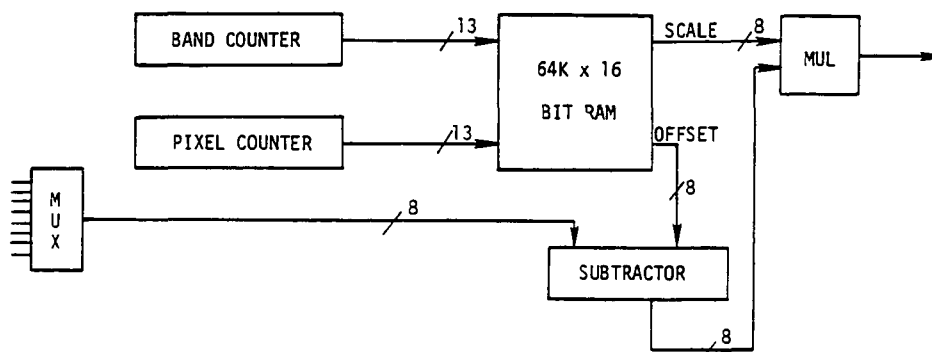


Figure 3-7 A Parallel Implementation of Radiometric Correction

The following times are assumed for this configuration

MUX settling time (not a problem since the RAM settling time is longer)	10 ns
Counter settling time	10 ns
SUB time	5 ns
RAM address stable to output stable	45 ns
MULTIPLIER settling	45 ns
	<hr/>
	105 ns

Worst Case Delay

This configuration is purely combinational logic and simply requires sufficient delay. By adding a fast latch between the RAM and multiplier, one could add a degree of pipelining and reduce the RAM speed requirements from 45 to 90 ns.

Sizing

An approximate size estimate can be computed for this configuration by using the following parameters. On the average, a memory cell implemented using a 1μ MOS technology requires a square 12μ on a side. Using that number, the required 128k byte RAM requires 1.5 cm^2 space on a chip.

Such high densities require polysilicon interconnects but the high resistivity of polysilicon, coupled with the extremely large chip distances to bring signals out, make it highly unlikely that this can be implemented on a single chip in the foreseeable future.

Alternative

An alternate design is shown below.

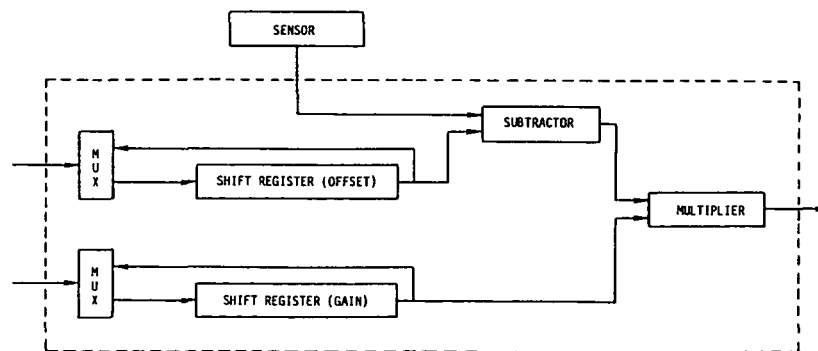


Figure 3-8 Radiometric Correction Using Shift Registers

Shift registers are proposed rather than RAMs since, in this configuration, they are more easily built, requiring only a MOS inverter and a pass transistor per cell. Memory is accomplished by storing charge on the gate of the inverter. Since the operation of the system will require essentially continuous clocking, refresh is automatic.

Such a design thus requires approximately 36,000 MOS transistors for the memory, plus the multiplier, which is within the feasible realm of VLSI, in addition, a response is now required every 800 ns rather than every 105 making the speed constraint more reasonable.

In VLSI, computation is cheap, communication is expensive, which is why this design places a separate multiplier on each chip rather than multiplexing inputs to a single multiplier off chip. This design provides simplicity and modularity, which would be more difficult to attain using a single multiplier.

3.6.2 A Microprogrammed Along-Scan Processor

AMD's 2900 series of bit slice bipolar microprocessors consist of a set of chips, some for computation, some for control, and some for condition code testing. In this design since we are building a special-purpose signal processor rather than a generalized computer, we will restrict ourselves to the use of two devices, the 2901 ALU chip, and the 2909 microprogram sequencer.

Attached to these two chips will be auxiliary chips required for signal processing, and chips required for microprogramming. These are shown in Figure 3-9.

Functionally, the system follows the design specified in Figure 3-8 of RTI's report to NASA [11]. The computation of memory addresses is taken over by the 2900 system, as well as calculation of distortions. As before, cubic convolution weights are stored in an 8 bit ROM (CCROM), and $A(I)$ values derived from distortion information is stored in 128 x 8 RAM (ARAM). Data enters a 4 x 8 file according to the 2 bit address input, and is written into the file at the time of the RFWE pulse.

Since addresses are provided to these memories by the 2900, they must be latched. Address latches internal to the memories are assumed, and new addresses will be latched for ARAM, CCROM or RF, at the time of ARAMADDR, CCROMADDR, or RFADDR pulses, respectively.

As before, accumulate functions occur at the time of the ACCP pulse.

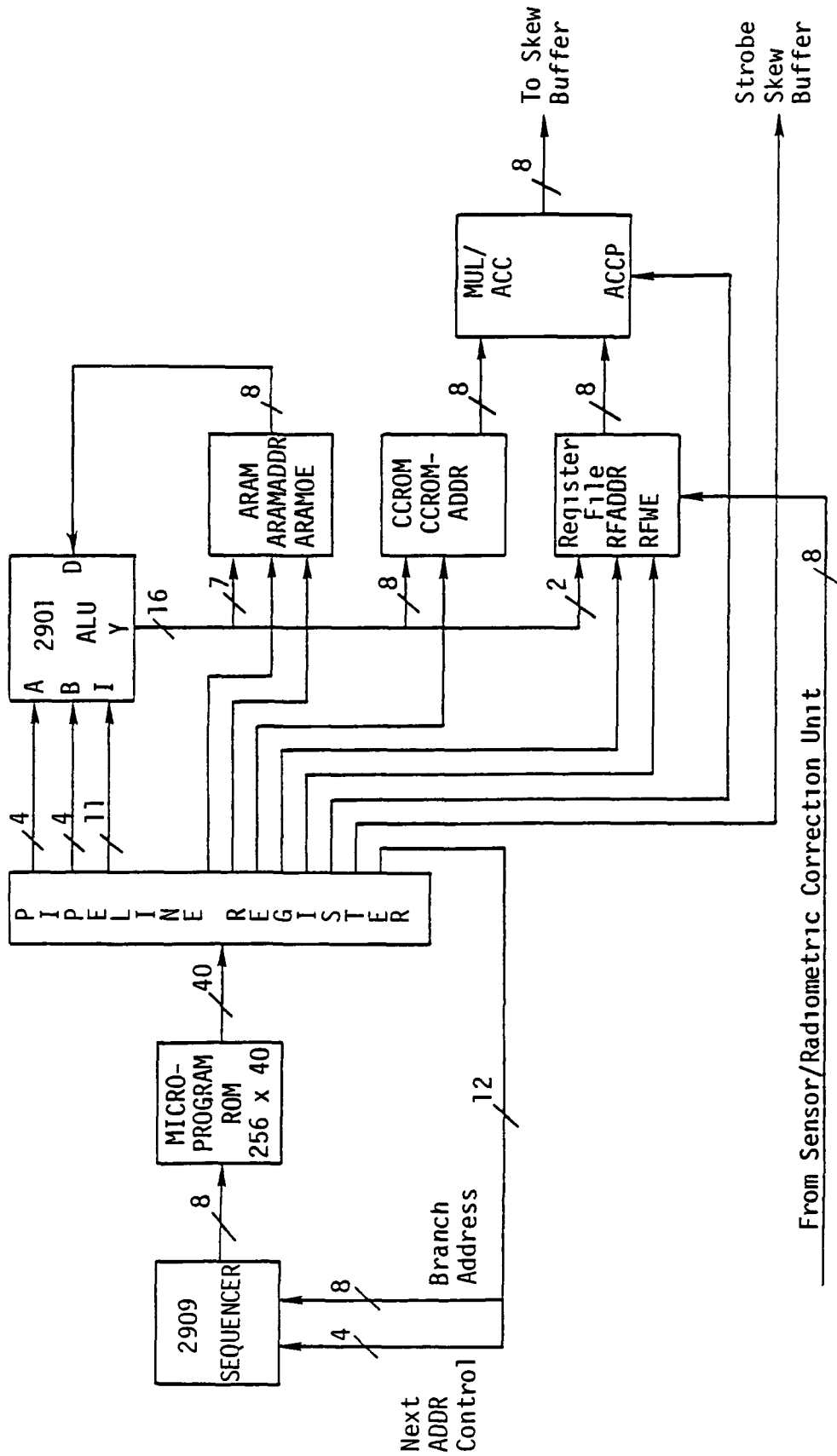


Figure 3-9 Along-Scan Processor Design

2901 Operation

Before describing the systems operation, it will be necessary briefly to describe the operation of the 2901 and 2909 chips. The 2901 block diagram is shown in Figure 3-10. It consists of an ALU, a 16 x 16 (in this configuration) register file, and an additional work register, designated "Q". The register file is dual ported, and one can access a register via either of two address, "A" or "B".

Input to the ALU can come from either the registers or from the D inputs.

2909 Operation (Figure 3-11)

The 2909 controls the next address to be presented to the microprogram ROM. The address may come from the microprogram counter register (part of the 2909), from the most significant 8 bits of the microprogram control word, in the case of a subroutine call or a jump microinstruction, or from the pushdown stack (also part of the 2909) in the case of a return from subroutine.

Along-Scan Microsequence

Register definitions

R0	Input pixel counter (IPC)
R1	Output pixel counter (OPC)
R2	SUM (to accumulate distortions)
R3	MASKFC00 (used in masking operations)
R4	Work register
R5	MASK 0003
R6	Work register
R7	Mask 000C

The register-transfer notation described below is in one-to-one correspondence with the microprogram in Table 3-4.

- 1 Clear accumulator, Increment input pixel counter accomplished by setting ADD of register 0 to zero, with carry-in bit set
- 2 Transfer OPC to output of 2901, accomplished by setting ALU ADD of register 1 to zero, with output to Y, and 2901 output enable set. Strobe ARAM to latch its address inputs and begin a memory cycle
- 3 Enable tri-state outputs of ARAM, read this value through 2901 D inputs, add it to register 2, results to register 2, call subroutine

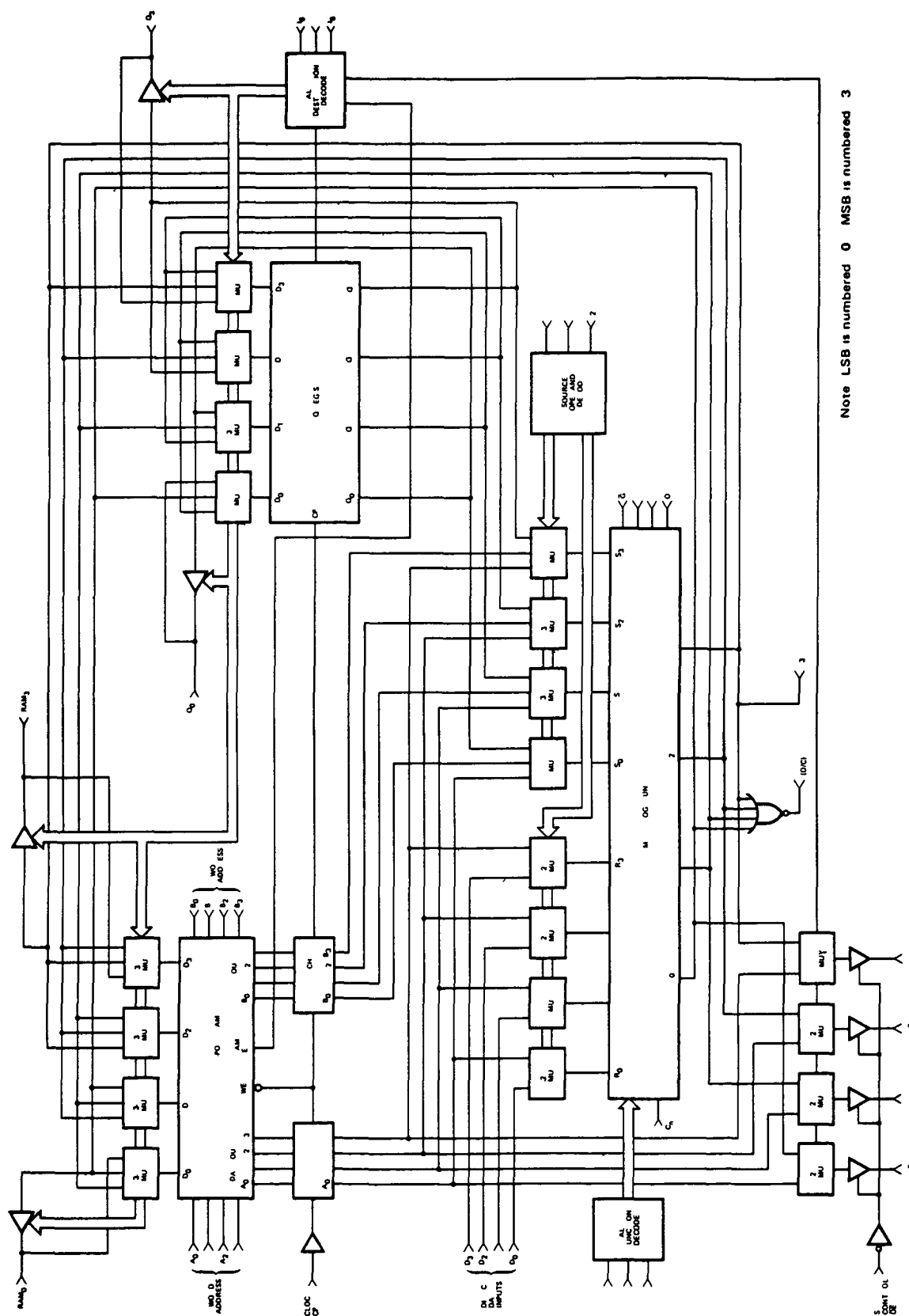
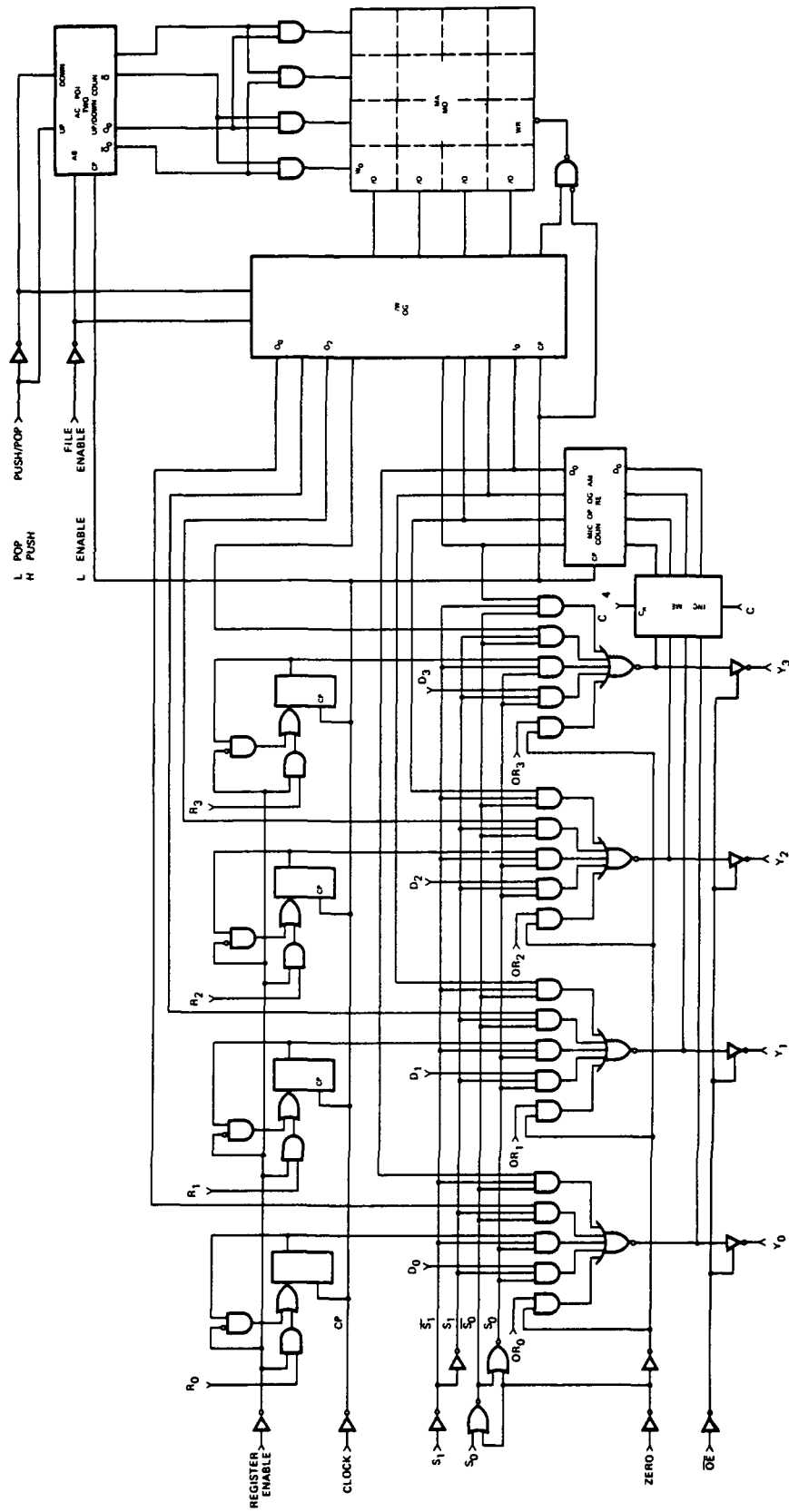


Figure 3-10 Detailed Am2901 Microprocessor Block Diagram [12]



Note R₁ and D₁ connected together on Am2911 and OR₁ removed

Figure 3-11 Microprogram Sequencer Block Diagram [12]

FE	PUP	S ₁	S ₀		
CALL	0	1	1	1	NEXT
RTS	0	0	1	0	ADDRESS
JMP	1	1	X	1	CONTROL

- 4-6 Increment input pixel counter, call subroutine
- 7 Increment OPC, Strobe data to skew buffer
- 8 Strobe the read-write file enable to read in a new pixel value,
go to step 1

Most operations are performed by a subroutine, which functions as follows

- 1 Use the mask in register 7, AND it with the SUM (R2), storing the result in the Q register
- 2 Transfer the Q register into register 4
- 3 Use the mask in register 5, AND it with IPC (R0), store the result in Q
- 4 Add Q to R4, output to CCR0M, forming the address to the ROM
- 5-6 Shift the IPC 2 bits left into R4
- 7 Add R4 to R6, output to form the new address to RF
- 8 Strobe the accumulate pulse, return from subroutine

A Variation

A good deal of the machine cycles used in this design are involved in the computation of the two bit address used by the input file. This file could be replaced by a simple 8 bit, four cell shift register, and a great deal of complexity removed. In this design, another bit in the microprogram word is provided to control shifting in of data. Two alternatives are then available using a single multiply/accumulate chip, and shift register, or using four multipliers and a summer. These alternatives will be discussed separately.

Variation 1 Using a shift register and a multiply/accumulate chip

Figure 3-12 shows one way to accomplish this function. Every four shifts, data is transferred to the skew buffer from the output of the MUL/ACC, and new input data is read into the top latch, via the MUX, replacing the oldest value.

Variation 2 Using a simple shift register and four multipliers

The MUX can be eliminated and control simplified by using a separate multiplier on each stage of the shift register. Now, CCR0M must be reorganized to permit 4 outputs to be simultaneously available, but the microprocessor speed demand is cut by 4. (See Figure 3-13)

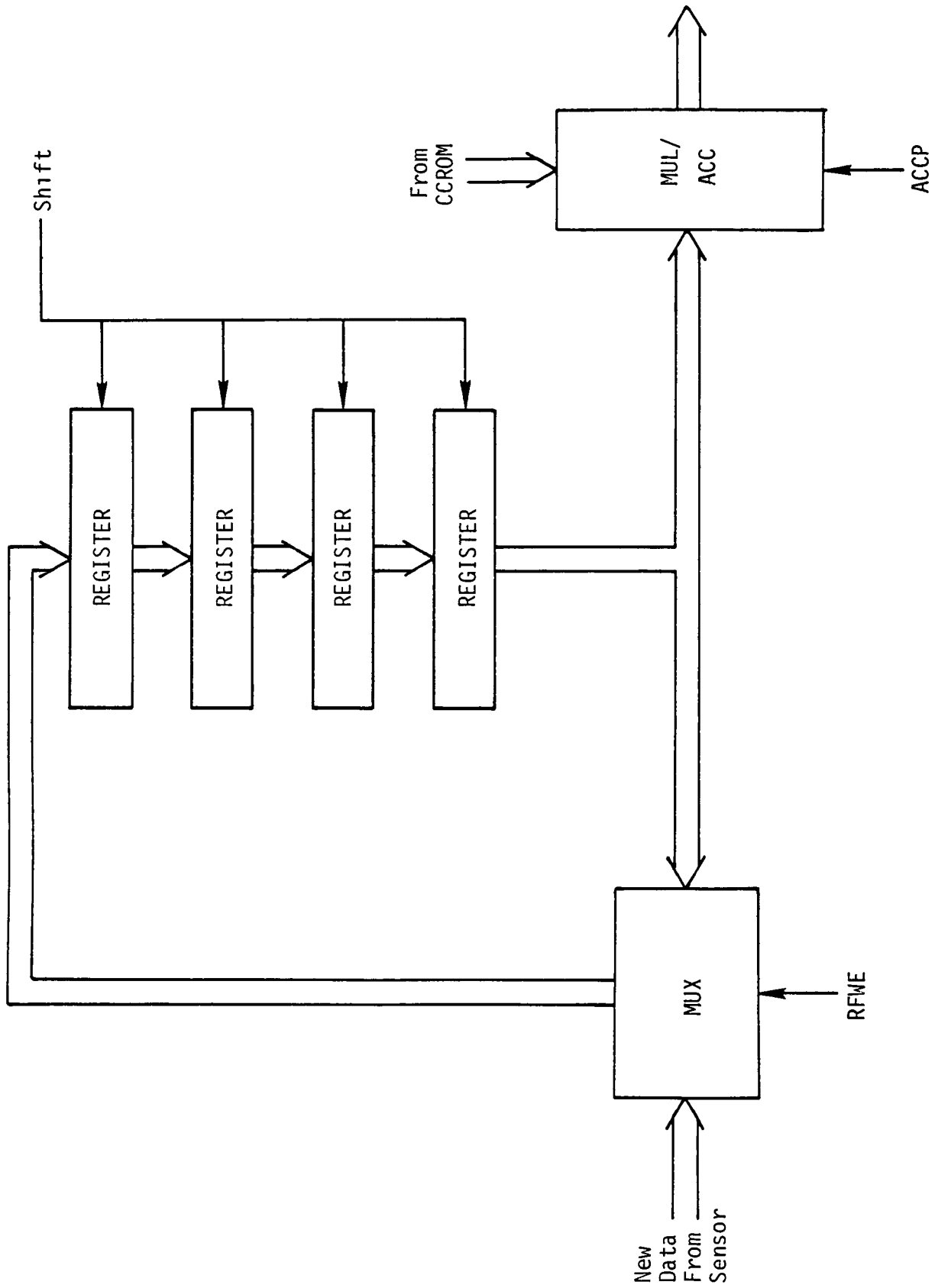


Figure 3-12 Alternative Input Register File Structure

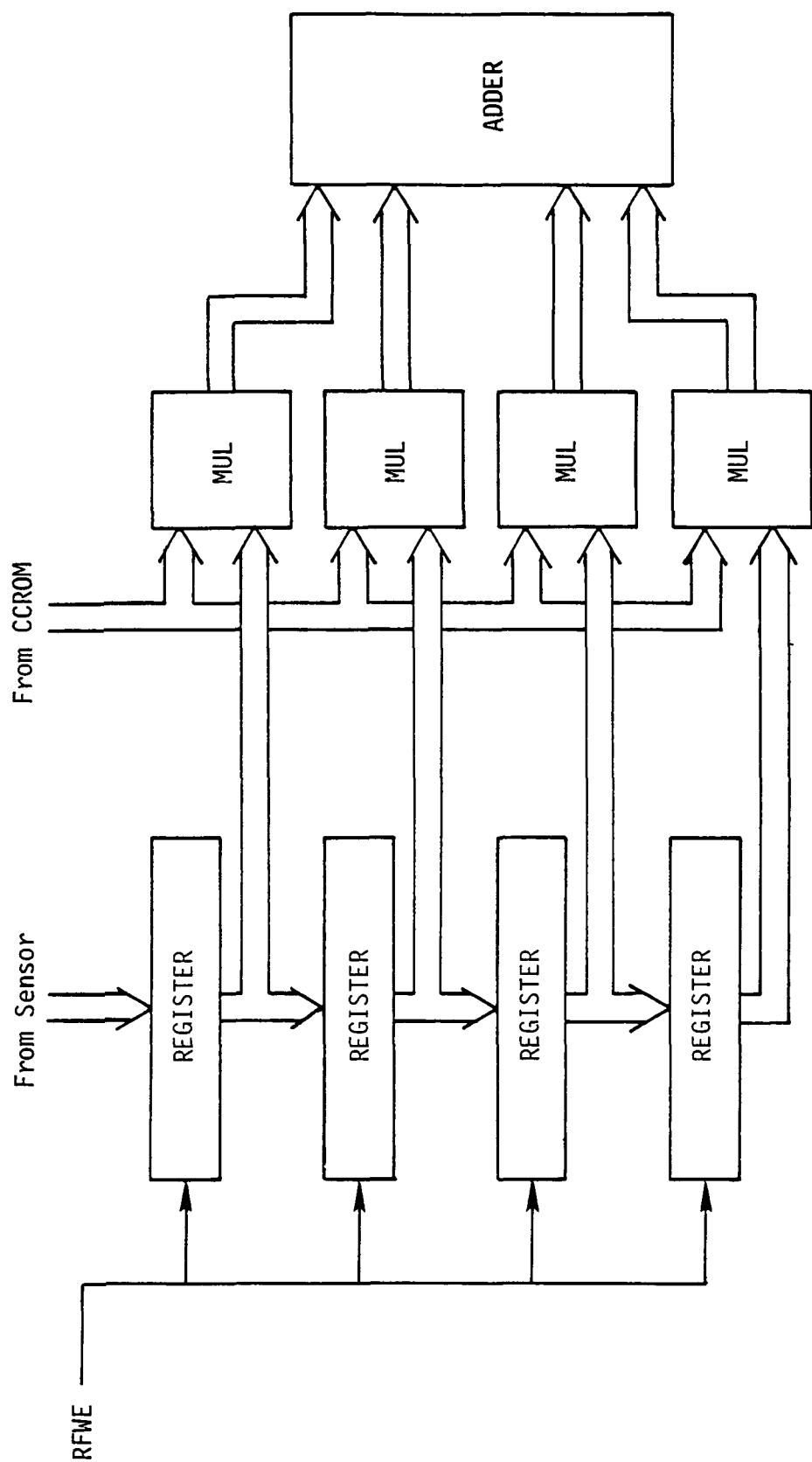


Figure 3-13 Faster Alternative Input Register and Arithmetic Section Design

Design Evaluation

These designs have been structured to incorporate only the most essential elements of geometric interpolation. The same designs could be used for along-scan and across-scan interpolation, assuming skew buffer address calculation is performed separately.

One function not detailed here is the ability to detect a distortion overflow. In this case, it is necessary to simply read another input pixel without outputting a pixel. This can be accomplished by detecting the overflow condition from the ALU after instruction 3, and conditionally jumping to another address.

A more serious deficiency in this design is that it does not detail a mechanism for allowing the general purpose computer access to the system for control, initialization, and loading of ARAM. This would be most appropriately handled by having the CPU set a bit requesting service, which would be tested during microinstruction 1, similar in operation to a conventional microprogrammed interrupt handler. Data or commands from the CPU could be read by the microprocessor via the D inputs to the ALU.

Sizing The first design shown, if implemented in current technology, would require

		<u>per chip</u>	
		<u>gates</u>	<u>devices</u>
2	2909	225	1,800*
4	2901 ALUs	545	4,350*
1	40 bit latch	200	1,600*
1	256 x 40 ROM (μ p ROM)		41,000*
1	128 x 8 RAM (ARAM)		6,000*
1	256 x 8 ROM (CCROM)		8,000*
1	4 x 8 RAM (RF)		200*
1	8 bit multiply/accumulate chip	404	5,944
Total devices-----			83,744

* estimated

The first design requires 32 cycles per interpolation, requiring, using the current 100 ns standard clock, 3.2 μ s. A new pixel is available every 600 ns, so a speed up of 5:1 over current technology would be needed.

Variation 1 trades an additional 32 bits of memory and a multiplexer for removal of microinstructions 5, 6, and 7 from the subroutine (with some modifications of other microinstructions), reducing the cycle time to 2 us

Variation 2 leaves the microprocessor with nothing to do except compute distortions, making 2900 implementation marginally feasible today. Alternative 2, however, adds approximately 1,200 gates, or 15,000 devices to the design, increasing die size, cost, and power consumption.

Remarks

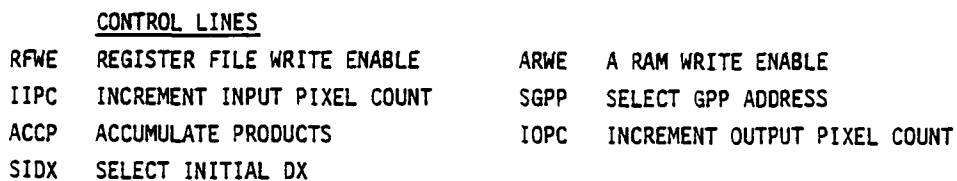
Computer-aided design techniques are expected to make it increasingly easy to produce custom VLSI. The module-based CAD techniques will enable the user to specify custom VLSI to be built from standard modules such as ALUs, shift registers, etc. We have shown in this section that it is reasonable to construct the interpolation subsystem from a VLSI package designed around existing parts, the 2900 series family. Whether such a design is the most appropriate use of VLSI technology is another question.

The most effective use of VLSI technology is to make as much use as possible of pipelining. The 2900 series of chips are designed particularly for the user who desires to build a high speed "general purpose" (in some sense) data processor with special instructions. It is not really well suited to particular high speed operations which are executed repetitively with little or no change in the flow of control through the operation.

The parallel structure of Figure 3-14, with some modifications, is well suited to VLSI implementation. RTI therefore recommends against adoption of the design given in this study as anything other than an interesting and educational exercise. However, use of the 2909 sequencer, or that type of device is the appropriate mechanism for implementation in VLSI of the control subsystem.

Furthermore, shift registers are easier to implement in VLSI than counter/RAM structures, so RTI proposes the design of Figure 3-14 with the register file and counter circuit replaced with a recirculating shift register. RTI feels this design is the most appropriate for VLSI implementation. In addition, a 2909-like microprogram sequencer would be used to control the data flow, handle "interrupts" when the CPU needed to write the ARAM, and provide control of shift/noshift in the case that 2 input pixels occur without an output pixel (overflow condition).

This design would make it reasonably easy to construct a single channel of the along-scan, or the across-scan interpolation systems on a single VLSI chip.



34

4 0 On-Board Signal Processing Functions

This section of the report will survey those functions applicable to signal processing on board spacecraft. Each function will be defined mathematically, but without discussion of the particulars of the mathematics.

Some typical applications of these functions are mentioned, as they might be used on spacecraft.

Various implementations of each function are described, with particular reference to the potential impact of the VHSIC program.

4 1 Convolution and Correlation

4 1 1 Definition

The convolution of two signals is expressed as

$$y(t) = \int_{-\infty}^{\infty} x(u)h(t-u)du,$$

Or, in sampled data systems, as

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k)$$

By reversing the time axis of one of the functions above, one arrives at a definition for correlation. The cross correlation $g(x)$ of two functions $a(x)$ and $b(x)$ is defined as

$$g(x) = \int_{-\infty}^{\infty} a^*(u)b(x+u)du,$$

or, in sampled data systems, as

$$g(n) = \sum_{k=-\infty}^{\infty} a^*(k)b(n+k),$$

where $*$ indicates the complex conjugate if the functions are complex.

The auto correlation of a function is likewise defined by

$$R(x) = \int_{-\infty}^{\infty} x^*(u)x(u+x)du \text{ or}$$

$$R(n) = \sum_{k=-\infty}^{\infty} x^*(k)x(k+n)$$

for discrete systems

4 1 2 Applications

On board space craft, the functions of correlation and convolution find principal application in signal detection Three typical instances are as follows

- a) Detection of the return pulse from a synthetic aperture radar system, typically using matched filtering techniques
- b) Detection of time shift in pseudo-random codes generated by GPS satellites to provide precise ranging information
- c) Identification and registration of ground control point features for precise image registration and determination of spacecraft position and attitude This application utilizes two dimensional correlations

4 1 3 Implementation of Correlation and Convolution

One can implement the functions of correlation or convolution in a wide variety of ways This variety can be partitioned into two distinct approaches, those which are centered around use of traditional computers and software, and those which utilize special purpose hardware

When considering computer-based methodologies, one is confronted with two options, to compute the correlation directly, that is, by taking a sum of products, or to compute it by transform techniques, as shown in Figure 4-1 The computational complexity of the Fast Fourier transform is $O(n \log n)$ where n is the number of points This can be considerably faster than the direct correlation, whose complexity is $O(n^2)$ However, this speed improvement depends on several factors, such as the number of points to be correlated or convolved whether the sequences are real or complex, and the type of FFT used For short sequences, less than about 100 points, the direct method is faster

4 1 3 1 Implementation of Correlation and Convolution Using Special-Purpose Hardware

Any device or system which implements the sum of products calculation can be used as a hardware implementation of such a function If, in addition, the data to be correlated can be stored in a shift register whose parallel outputs are directly connected to multipliers, the implementation is most straightforward

Figure 4-2 shows a direct implementation of the correlation function For

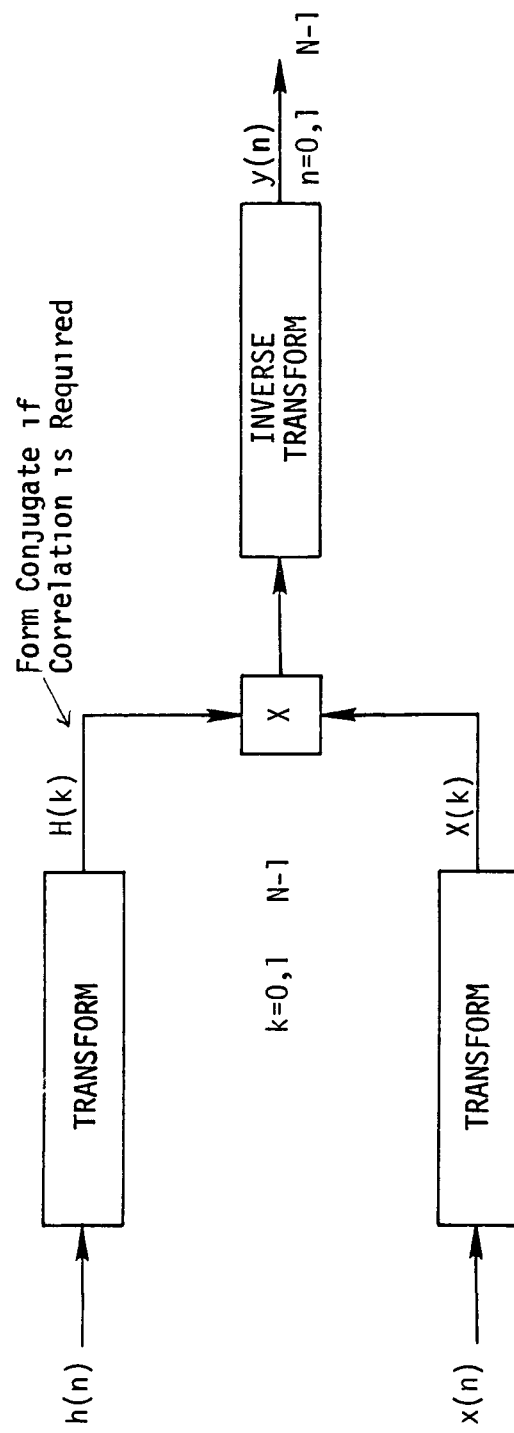


Figure 4-1 Transform Convolution

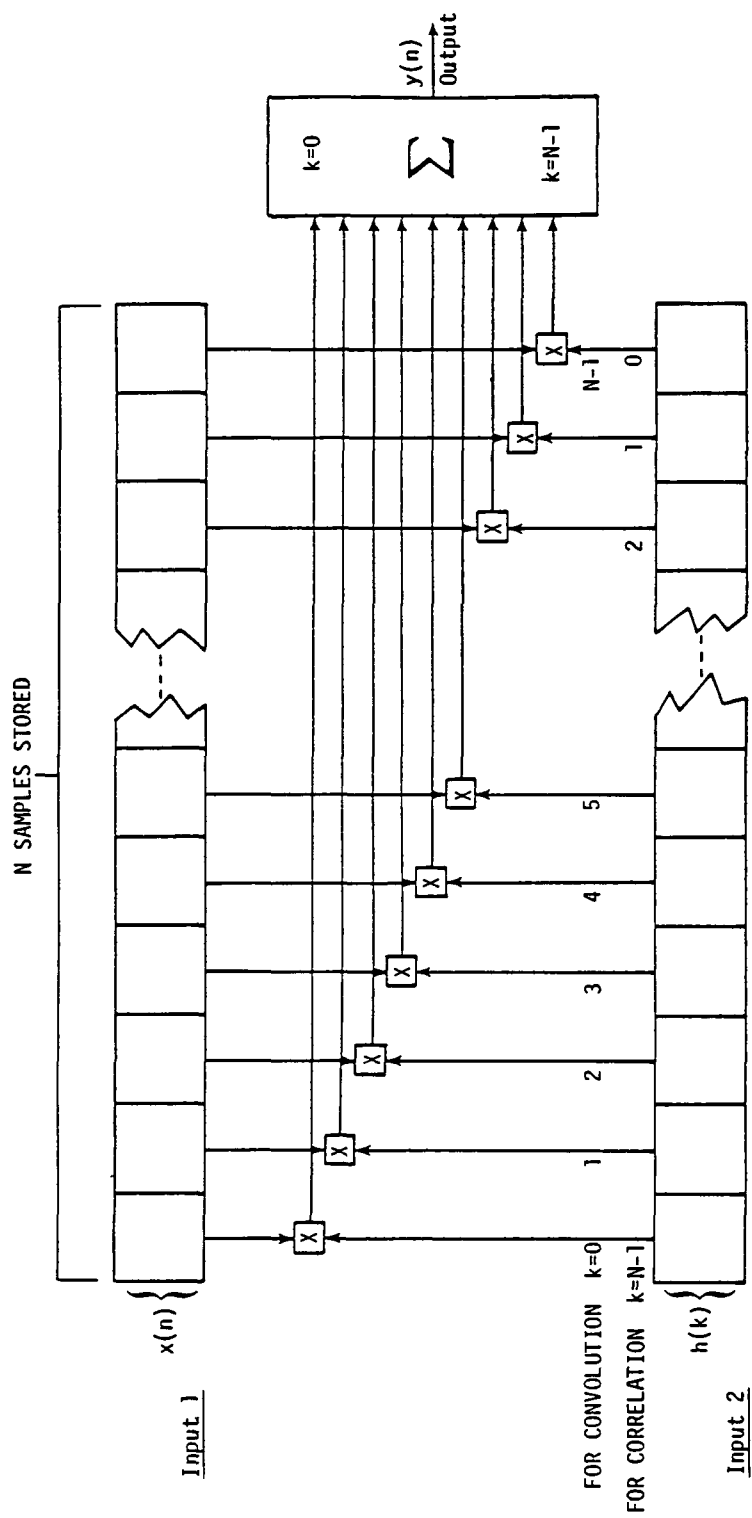


Figure 4-2 General Realization of Convolution & Correlation [13]

an n cell correlator, n multipliers are needed. Because of this large number of multipliers, a direct implementation in digital hardware is quite consumptive of power and chip space. An alternative digital implementation requires a single multiplier/adder/accumulator and rotates the data past. This option trades speed for hardware.

Transform correlation/convolution can also be performed using special purpose hardware. Section 4.2 of this report describes Fourier Transform pipeline structures. The pipelined FFT requires only $O(\log_2 n)$ components, however, there are some corresponding disadvantages to the use of FFT pipelines in this context. First, the nature of the FFT structure requires that the number of elements to be transformed be known in advance, and be a power of 2. Furthermore, the FFT requires much more sophisticated control hardware than direct correlation techniques. Direct hardware can be used to compute correlation shorter than those for which the system was designed, simply by filling with zeroes. Transform systems however cannot simply be filled with zeroes to accomplish correlation with fewer samples.

An alternative to direct digital implementations is provided by sampled analog processing using charge transfer devices. NASA has investigated such systems at length with the following conclusions:

- 1) If one of the signals being correlated is a constant, as would be time in many matched filtering applications, with constant impulse response, then split-electrode CCDs can provide clock rates of 5-10 MHz and third decades of dynamic range. Such units have been built which correlate signals up to 500 samples in length.
- 2) With variable tap weight devices, which allow correlation of two time signals, analog multipliers restrict the dynamic range to about 8 bit maximum. No analog/analog correlators have been built more than 64 cells long, with maximum functional clock rates of 500 KHz.
- 3) Using a digital shift register for one signal and a CCD for the other, Texas Instruments has produced a 16 cell analog/binary correlator which provides added accuracy with a 2 MHz sampling rate. Experimental results indicate the dynamic range available from analog-analog correlators is 7-8 bits, whereas analog-binary can provide 8-9 bits.

4 1 3 2 Surface Acoustic Wave Devices

For high frequency operations, surface acoustic wave systems can provide good performance. Figure 4-3 shows a plate convolver. The voltage V_a and V_b simultaneously launch two surface waves towards the center of the device. The duration and synchronization to the input pulses must be such that their interaction occurs entirely under the plate.

The interaction of the two acoustic waves produce a voltage on the plate electrode equal to the integral of the product of the waveforms.

Such a unit has a useful dynamic range of 60dB and time bandwidth products of over 1000. The major difficulty of SAW correlators is the insertion loss, up to 95dB for the plate convolver described, and as high as 30dB for other SAW correlator designs.

4 1 3 3 The Impact of VHSIC on Correlation Operations

It is an interesting exercise to compute the maximum speed with which correlation can be performed, assuming a digital implementation. The structure of Figure 4-2 is used for the calculation, with a separate digital multiplier at each stage. Addition can be performed by a binary tree of adders.

For a correlation n cells long, n multipliers and $2n$ adders will be needed. Assuming the maximum possible speed, 2 gate delays for multiplication and addition, since the signal must pass through $\log_2 n$ adders, we find propagation delay of $(2 + 2 \log_2 n) \tau$ seconds is required. If τ , the propagation delay per gate, is 10ps (a realistic upper bound for foreseeable technology, even allowing Josephson Junction devices), a 1024 cell unit will still require 220ps to settle and cannot be operated synchronously at higher than 4 gigahertz.

In summary, silicon-based implementations of correlation can be organized by the following table:

Traditional (i.e., CPU-based) Implementations -

"Short" correlation (100 points or less) - simple sum of products is best - n^2 operations

"Long" correlations - use of FFT requires $O(n \log n)$ operations

Special Hardware -

Fully digital sum of products-simple, high speed, accurate, massive power consumption

Digital Sum of Products Using a Single Time-Shared Multiplier -
more complexity, lower speed, much less power

Digital FFT - a variety of pipelined approaches available- described in Section 4.2

Analog Correlators -

If the signal to be correlated is known in advance, fixed tap weight CCD filters can be used [14], providing reasonably high dynamic range (about 10 bits of accuracy, maximum), high speed (10 MHz max clock rate), and very low power

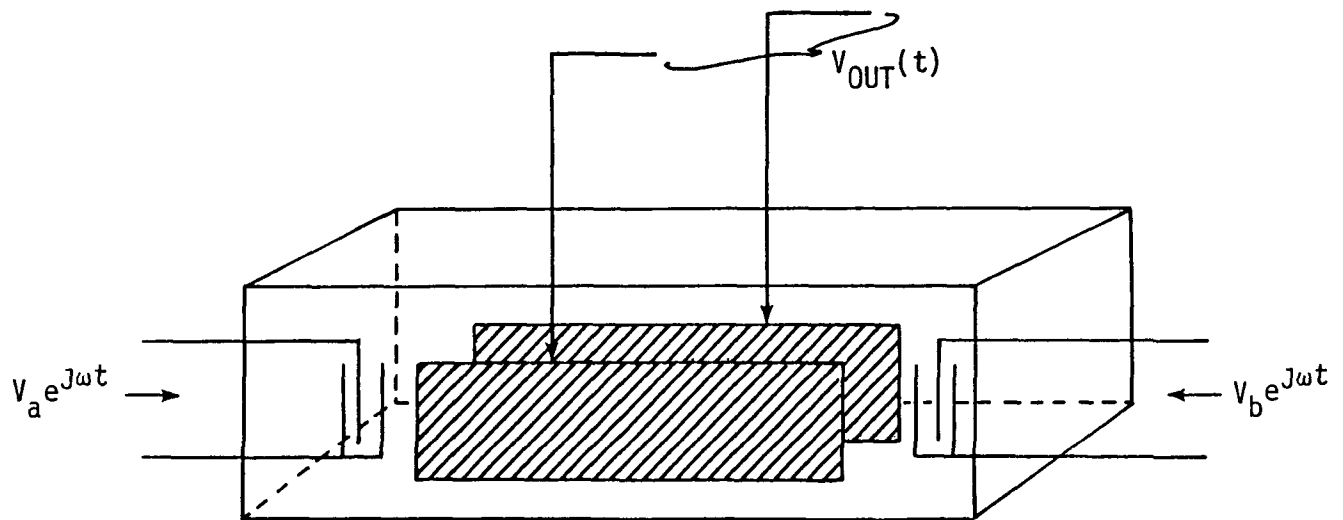


Figure 4-3 Plate Convolver [13]

4 2 Transforms

4 2 1 Discrete Fourier Transform (DFT)

The DFT of a sequence of N values is defined to be

$$F(k) = \sum_{n=0}^{N-1} f(n)e^{-j(2\pi/N)nk}, \quad k=0,1, \dots, N-1$$

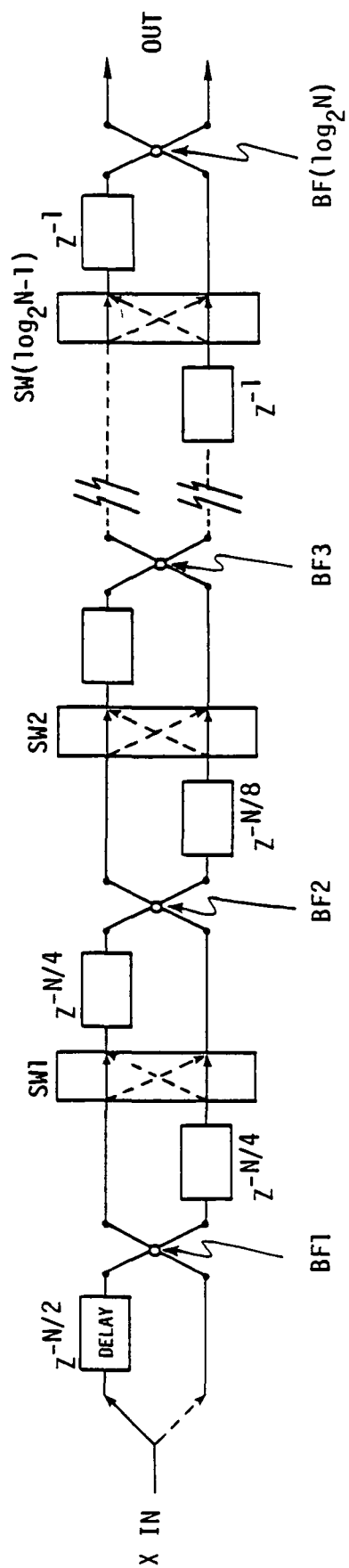
where $f(n)$ is the original sequence and $F(k)$ is the transformed sequence. The DFT can be used to implement fast convolution and correlation as discussed previously.

4 2 1 1 Fast Fourier Transform (FFT)

The FFT is a well-known algorithm for reducing the number of arithmetic operations required to compute the DFT. The details of the algorithm are readily available in the literature and will not be repeated here. The significant result of using the FFT over directly computing the DFT is a reduction of the number of complex multiplications from $O(N^2)$ to $O(N \log_r N)$, where N is the number of samples and r is the radix of the FFT, such that $N=r^M$ for some integer M . The most commonly used radices are 2 and 4. When the sequence lengths can be chosen to be a power of 4, the radix 4 FFT appears to be the best choice as a trade-off between speed and hardware complexity [13]. Radix 2 results in simpler hardware which is fast enough for many applications, and allows a more flexible choice of sequence lengths.

4 2 1 2 Special-Purpose Hardware

Figure 4-4 shows a pipelined implementation of a radix 2 FFT. The system will contain M stages, $M = \log_2 N$, with each stage containing delay elements (shift registers), a digital "switch" (gating logic), and an arithmetic element to implement the FFT butterfly operation. The butterfly computation involves a sum, a difference, and one complex multiplication. The data rate available from such an implementation is determined by the speed of the slowest element, usually the multiplier. When input buffering is used, allowing the pipeline to run at 100% efficiency, then the data rate can be as high as the maximum multiplier rate.



BF = Butterfly
SW = Switch

Figure 4-4 Radix 2 Pipeline FFT [13]

times the radix. This assumes a pipelined implementation of the butterfly itself, with latches between sequential arithmetic elements. Thus the maximum clock period is equal to the multiplier delay plus one latch delay, or about 125 ns with current technology for a 16-bit implementation, resulting in a data rate of 8 MHz.

The basic subsystem used in building pipelined FFT hardware is the arithmetic element used to compute the butterfly operation. The radix 2 butterfly element requires four real multipliers and 6 real adders. Table 4-1 derives an estimate of the number of devices (transistors) required to implement a 16-bit butterfly element. Such a subsystem could be fabricated on a single chip with state-of-the-art LSI technology. However, with the submicron VLSI technology anticipated for the 1985 timeframe, chips with one order of magnitude higher device density will be feasible. To illustrate the significance of this capability for FFT hardware, let us consider the complexity of the radix 4 butterfly subsystem, which allows a doubling of the data rate, and a reduction of the number of stages by half. Table 4-1 derives the number of devices required by a 16 bit radix 4 butterfly element. This subsystem would easily fit on a single VLSI chip using submicron technology. Use of such a component in pipelined FFT hardware instead of the radix 2 butterfly chip currently feasible would increase the data rate by at least a factor of 5 and greatly reduce the component count, power consumption and circuit complexity.

One problem with implementing a radix 4 butterfly on a single chip is getting the data in and out. The requirement is for 4 complex inputs, 4 complex outputs, and 3 complex "twiddle factor" inputs. It might be desirable to store the twiddle factors in a ROM (read-only memory) on the chip. The requirement is for $2N \times b$ bits of ROM, where N is the maximum transform length, and b the word size in bits. For $N = 1024$ and $b = 16$, the ROM size is 32K bits, which can be placed on the same radix 4 butterfly chip using submicron technology. The other 8 complex inputs/outputs will have to be multiplexed to achieve a reasonable pinout configuration.

Another illustration of the capabilities of VLSI in transform hardware is illustrated in Table 4-2, which shows that a pipelined 256-sample 16-bit FFT can be built with under 500,000 devices (transistors). With submicron technology, this device can be fabricated on a single chip with a data rate of 32 MHz and power consumption of around 3 watts.

Table 4-1 16-Bit Butterfly Requirements

	RADIX 2	RADIX 4
number of real multipliers	4	12
x 16,600 devices=	66,400	199,200
number of real adders	6	22
x 400 devices=	2,400	8,800
number of devices for control and timing logic	1,200	2,000
Total devices	<u>70,000</u>	<u>210,000</u>

Table 4-2 Radix 2, 16-Bit, 256-Sample Pipelined FFT Requirements

	Number Required	Devices Each	Total Devices
Full Butterfly Elements	6	70,000	420,000
Additional Adders	12	400	4,800
ROM (bits)	256 x 16	3	12,288
Shift Register Cells	768 x 16	3	36,864
Digital Switches	224	4	896
Nand Gates	640	3	1,920
Other Control Logic Gates	256	4	1,024
Total Devices			<u>478,000</u>

4 2 2 Other Transforms

On-board signal processing functions may utilize less common transforms such as Walsh, Haar, Chirp-Z, Prime, Discrete Cosine, Karhunen-Loeve, and Number Theoretic Transforms. Many of these can be implemented with fast algorithms similar in structure to the FFT. In any case, special-purpose hardware to implement these transforms will benefit from similar improvements in performance from VLSI as those detailed for the FFT in section 4 2 1 2. Details of these transforms and some hardware implementations can be found in the study by Ferranti [13].

5 0 VHSIC Availability

The types of VLSI parts which are currently available and which can be anticipated to be available in the near future have been surveyed in the Technology Forecast section of this report (Section 3 0), and those which will be available in the more distant future are discussed generically in the Beyond VHSIC section (section 6 0) As far as what particular devices will result from the current VHSIC program, the reports of the VHSIC phase zero contractors stand as the only source

Unfortunately, these reports are company confidential at this instant of time, and therefore not available The phase zero reports are due to be delivered to the VHSIC office, in the Office of the Undersecretary of Defense, on January 1, 1981 They can be expected to become public domain in three or four months In the meantime, they may be available to government offices

6 0 Beyond VHSIC

In a 1979 paper [15], R W Keyes of IBM's T J Watson Research Center extrapolated the technological trends to the end of the century. He forecast the integration level (Figure 6-1), power-delay product (Figure 6-2), and cooling capability (Figure 6-3), to predict the progress of packaged logic delay (Figure 6-4)

By comparing Keyes' figures for 1980 with technology known to exist, his figures can be seen to be slightly conservative in the sense that faster chips exist now, than those predicted by the graph. However, his predictions are for delivery of systems utilizing the specified technology

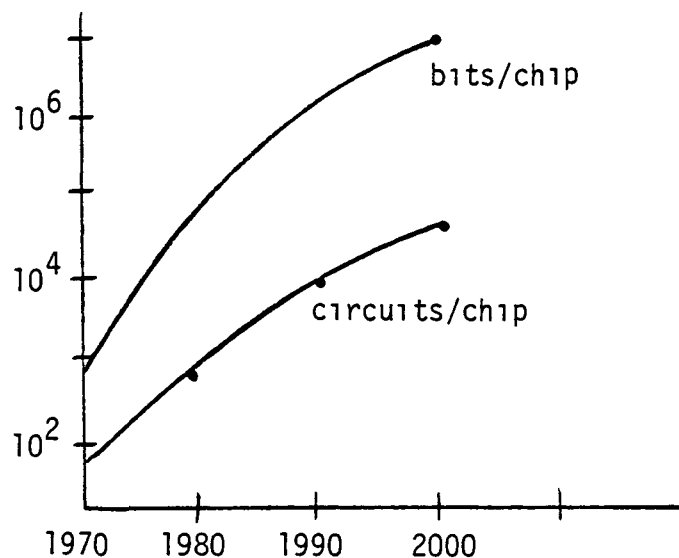


Figure 6-1 Levels of Integration [15]

Another mechanism for predicting the future state of commercially available VLSI systems is to look at what the currently most sophisticated laboratory products are. Two examples will be considered, one from silicon and one from gallium arsenide

6 1 Silicon

Bell Telephone Laboratories announced [16] in December 1980 the successful fabrication and testing of silicon MOS devices with 0.3 to 0.4 μm channel lengths. These devices have been tested in ring oscillator configurations and found to

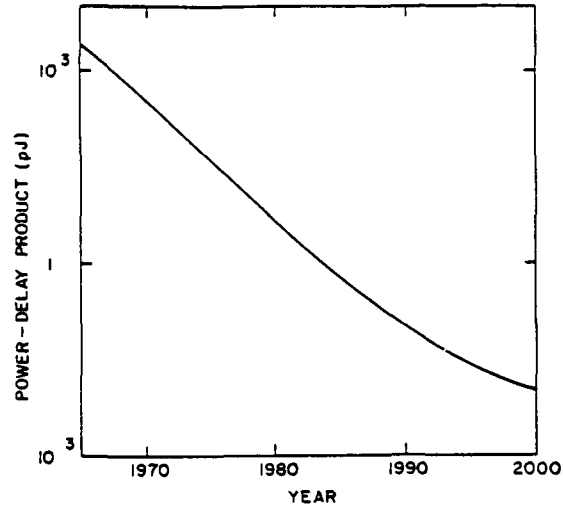


Figure 6-2

The power-delay product of logic, including early vacuum tube as well as transistorized computers. The extrapolation is drawn to approach a limit of 0.01 pJ [15]

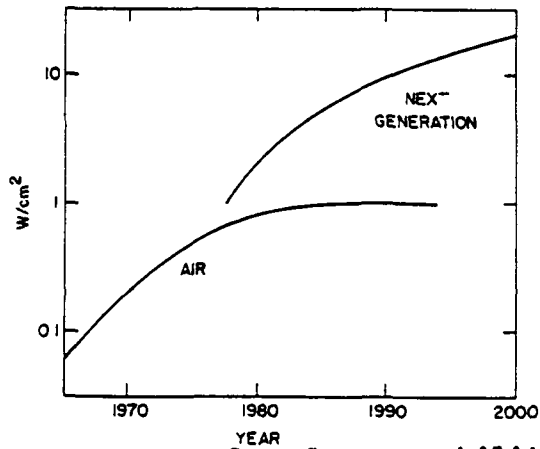


Figure 6-3

A projection of cooling capabilities for logic chip packages. Forced-air cooling is probably limited to about 1 W/cm². New cooling technologies, in which heat is transferred to a liquid without the intermediary of air, are emerging [18], [19], and it has been assumed that they will extend cooling capability to 20 W/cm² [15]

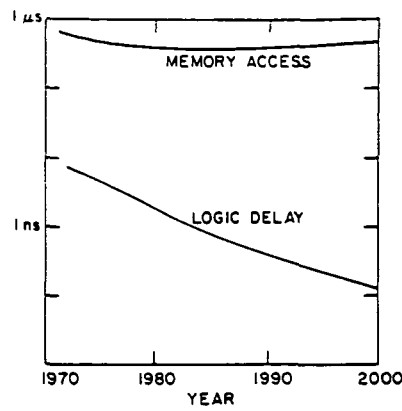


Figure 6-4 Packaged logic delays and main-memory access times calculated from models and extrapolations of technology [15]

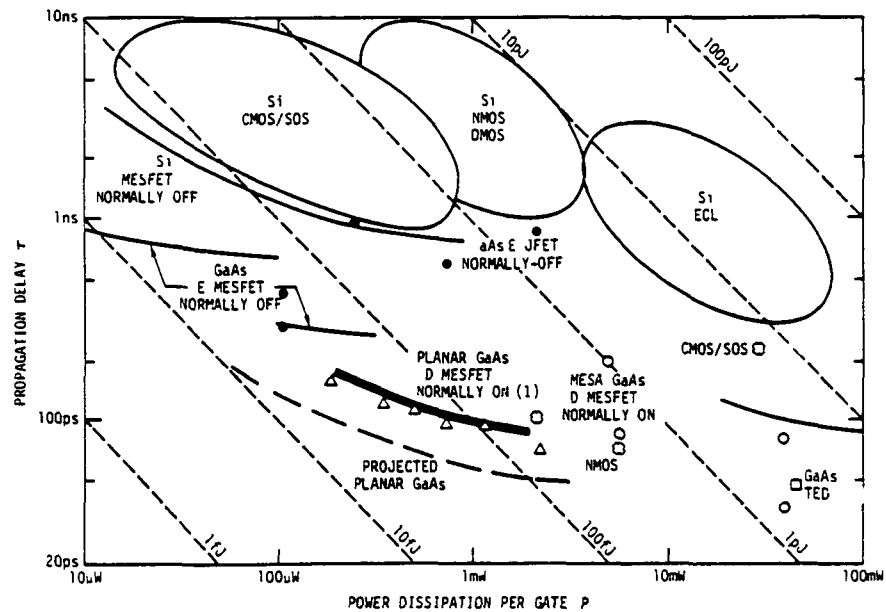


Figure 6-5 Map of propagation delay versus power dissipation per gate comparing published results for GaAs and Si IC technologies [17]

have switching speeds of 40 ps/gate. The testing of a chip in ring oscillator configuration can be misleading, as R C Eden [17] points out. "In typical silicon IC's (NMOS for example), there is about an order of magnitude difference between small inverter ring oscillator speeds and the speeds in real circuits fabricated from the same technology. About half of this speed loss results from the fan-out loading in real circuits and the rest comes from parasitic substrate capacitances incurred by use of a conductive substrate."

In a telephone conversation with Dr. Al Zacharias, of Bell Labs, RTI learned that his tests were performed with reasonable, multi-gate loading on the devices, and "probes hanging directly off the chip". One could thus conservatively estimate that this class of devices could operate with 100 ps/gate delays.

Straightforward extrapolations of technology typically require 2-3 years to move from the laboratory to the production line. The Bell Labs chips, however, are fabricated using X-ray photolithography, a new technology, rather than a simple extrapolation. There do not exist production line X-ray lithography systems today. Five years is probably a more reasonable time frame in which to expect this technology to mature.

6.2 Gallium Arsenide

The technology which most competes with silicon VLSI for future applications is offered by gallium arsenide devices.

Gallium arsenide has an electron mobility which is five to six times that of silicon, and therefore, under low field conditions, a GaAs device could be expected to be 5-6 times as fast as silicon. However, use of electron mobility to compare speeds can be misleading. Under high field conditions, electrons rapidly achieve saturation velocity, a velocity which is comparable in both silicon and gallium arsenide.

In practical switching applications, the devices are operating in high field, saturated velocity conditions only part of the time, and in transient field conditions a significant part of the time also. Under transient conditions, the higher mobility provides its advantage. These two factors combine to provide an expected intrinsic speed advantage of GaAs over silicon of 2 to 3 times.

Researchers in compound semiconductors have been experimenting with submicron lithography longer than silicon researchers, and consequently the submicron technologies are further advanced in the compound semiconductor areas. This explains,

to some extent, the rather impressive speed results which have been published to data. Figure 6-5 shows these results. As silicon device sizes become smaller as reported earlier in this section for example, the differences become less extreme.

However, the low capacitive loading from the semi-insulating GaAs substrate and higher mobility of electrons combine to produce extremely high functional throughput rates. Eden [17] has predicted that if the achievement of very complex ultrahigh-speed GaAs VLSI circuits proves possible a functional throughput of 2×10^{14} gate-hertz could be achieved, twenty times the ultimate objective of VHSIC.

REFERENCES

- 1 Sumney, Larry W , "Pipelining Innovation Into VHSIC The Phase III Concept", Military Electronics/Countermeasures, May 1980
- 2 Sumney, Larry W , "VLSI With a Vengeance", IEEE Spectrum, April 1980
- 3 Mead, Carver and Lynn Conway, Introduction to VLSI Systems, Addison-Wesley Publishing Company, 1980
- 4 Hoeneisen, B , and C Mead, "Fundamental Limitations in Micro-electronics-I, MOS Technology", Solid State Electronics, Volume 15, 1972
- 5 Ferranti, LTD Digital Systems Division, "On-Board Signal Processing, Technology and Systems, Volume 2", Report No 3134, March 1978
- 6 Leonard, Milton, "Tough Choices Ahead For VLSI Chips", High Technology, May 1980
- 7 Clary, J B , "An Overview of the Impact of Technology on Military Maintenance Concepts", IEEE Microprocessor Workshop, John Hopkins University, January 1980
- 8 Aviziens, A , "Fault-Tolerant Computing—Progress, Problems and Prospects", Information Processing 77, Proc IFIP Congress 1977, Toronto, Canada, August 1977
- 9 Burchly, D D , L W Kern and W A Strum, "Specification of the Fault-Tolerant Space-Borne Computer (FTSC)," FTSC-6 Proc Sixth Annual International Symposium on Fault-Tolerant Computing, Pittsburgh, Penn , June 1976
- 10 Holmes-Siedle, A , "The Impact of the Space Radiation Environment on the Semiconductor Technology Used in Spacecraft Processor Equipment", General Technology Systems Limited, Report No GT 760653, July 1977
- 11 Ruedger, W H , Daluge, D R , Aanstoos, J V , "Concepts For On-Board Satellite Image Registration", NASA CR-159287
- 12 Advanced Micro Devices, Inc , The Am 2900 Family Data Book, AMD, Inc , 1979
- 13 Ferranti LTD, Digital Systems Division, "On Board Signal Processing, Technology and Systems Volume 1, Signal Processing Operations", Report No 3066, December 1977
- 14 Buss, D D , C R Hewes, and M DeWitt, "Charge Coupled Devices for Analogue Signal Processing", IEEE Conference Publication No 144, September 1976

- 15 Keyes, Robert "The Evolution of Digital Electronics Toward VLSI", IEEE Journal of Solid State Circuits, April 1979
- 16 Mhatre, G , "Bell X-Ray System Could Revolutionize IC Lithography", EE Times, December 1980
- 17 Eden, Richard C et al, "The Prospects for Ultrahigh-Speed VLSI GaAs Digital Logic", IEEE Journal of Solid-State Circuits, April 1979
- 18 Wilson, E A , "True Liquid Cooling of Computers", 1977 National Computer Conference Proceedings (Monvale, N J AFIPS Press, 1977), pp 341-348
- 19 Russel, R M , "The Cray-1 Computer System", Communications of the ACM, Volume 21, pp 63-72, January 1978

1 Report No NASA CR-165735		2 Government Accession No		3 Recipient's Catalog No	
4 Title and Subtitle Concepts for On-Board Satellite Image Registration Volume Three Impact of VLSI/VHSIC on Satellite On-Board Signal Processing				5 Report Date June 1981	
				6 Performing Organization Code	
7 Author(s) J V Aanstoos W H Ruedger W E Snyder				8 Performing Organization Report No RTI/1796/00-03F	
				10 Work Unit No	
9 Performing Organization Name and Address Research Triangle Institute Post Office Box 12194 Research Triangle Park, NC 27709				11 Contract or Grant No NASI-15768	
				13 Type of Report and Period Covered Contractor Report May 1980 to December 1980	
12 Sponsoring Agency Name and Address National Aeronautics and Space Administration Langley Research Center Hampton, Virginia 23665				14 Sponsoring Agency Code	
15 Supplementary Notes Langley Contract Monitor - W L Kelly, IV Final Report					
16 Abstract This report describes anticipated major advances in integrated circuit technology in the near future, and the impact they will have on satellite on-board signal processing systems VLSI (Very Large Scale Integration) will achieve dramatic improvements in chip density, speed, power consumption, and system reliability Such improvements will enable more intelligence to be placed on remote sensing platforms in space, meeting the goals of NASA's IAS (Information Adaptive System) concept, a major component of the NEEDS (NASA End-to-End Data System) program A forecast of VLSI technological advances is presented, including a brief description of the Department of Defense VHSIC (Very High Speed Integrated Circuit) program, a seven-year research and development effort begun in 1979					
17 Key Words (Suggested by Author(s)) On-Board Signal Processing VLSI VHSIC			18 Distribution Statement		
19 Security Classif (of this report) Unclassified	20 Security Classif (of this page) Unclassified	21 No of Pages 59	22 Price		